



Amazon EMR リリース ガイド

# Amazon EMR



# Amazon EMR: Amazon EMR リリース ガイド

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon の商標とトレードドレスは、Amazon 以外の製品またはサービスとの関連において、顧客に混乱を招いたり、Amazon の名誉または信用を毀損するような方法で使用することはできません。Amazon が所有していない他のすべての商標は、それぞれの所有者の所有物であり、Amazon と提携、接続、または後援されている場合とされていない場合があります。

# Table of Contents

Amazon EMR リリースについて .....	1
Amazon EMR 7.x リリースバージョン .....	2
Amazon EMR 7.x リリースのアプリケーションバージョン .....	3
emr-7.1.0 .....	3
emr-7.0.0 .....	28
Amazon EMR 6.x リリースバージョン .....	64
Amazon EMR 6.x リリースのアプリケーションバージョン .....	65
emr-6.15.0 .....	65
emr-6.14.0 .....	111
emr-6.13.0 .....	157
emr-6.12.0 .....	211
emr-6.11.1 .....	273
emr-6.11.0 .....	329
emr-6.10.1 .....	366
emr-6.10.0 .....	421
emr-6.9.1 .....	461
emr-6.9.0 .....	514
emr-6.8.1 .....	559
emr-6.8.0 .....	612
emr-6.7.0 .....	657
emr-6.6.0 .....	715
emr-6.5.0 .....	775
emr-6.4.0 .....	801
emr-6.3.1 .....	833
emr-6.3.0 .....	859
emr-6.2.1 .....	891
emr-6.2.0 .....	916
emr-6.1.1 .....	948
emr-6.1.0 .....	969
emr-6.0.1 .....	994
emr-6.0.0 .....	1011
Amazon EMR 5.x リリースバージョン .....	1032
Amazon EMR 5.x リリースのアプリケーションバージョン .....	1036
emr-5.36.2 .....	1036

emr-5.36.1 .....	1060
emr-5.36.0 .....	1103
emr-5.35.0 .....	1141
emr-5.34.0 .....	1167
emr-5.33.1 .....	1191
emr-5.33.0 .....	1219
emr-5.32.1 .....	1241
emr-5.32.0 .....	1263
emr-5.31.1 .....	1290
emr-5.31.0 .....	1309
emr-5.30.2 .....	1331
emr-5.30.1 .....	1350
emr-5.30.0 .....	1372
emr-5.29.0 .....	1394
emr-5.28.1 .....	1413
emr-5.28.0 .....	1431
emr-5.27.1 .....	1450
emr-5.27.0 .....	1467
emr-5.26.0 .....	1486
emr-5.25.0 .....	1505
emr-5.24.1 .....	1524
emr-5.24.0 .....	1541
emr-5.23.1 .....	1559
emr-5.23.0 .....	1575
emr-5.22.0 .....	1594
emr-5.21.2 .....	1613
emr-5.21.1 .....	1629
emr-5.21.0 .....	1646
emr-5.20.1 .....	1664
emr-5.20.0 .....	1680
emr-5.19.1 .....	1700
emr-5.19.0 .....	1716
emr-5.18.1 .....	1733
emr-5.18.0 .....	1749
emr-5.17.2 .....	1766
emr-5.17.1 .....	1782



emr-5.17.0 .....	1797
emr-5.16.1 .....	1814
emr-5.16.0 .....	1830
emr-5.15.1 .....	1847
emr-5.15.0 .....	1862
emr-5.14.2 .....	1879
emr-5.14.1 .....	1894
emr-5.14.0 .....	1910
emr-5.13.1 .....	1927
emr-5.13.0 .....	1943
emr-5.12.3 .....	1958
emr-5.12.2 .....	1973
emr-5.12.1 .....	1988
emr-5.12.0 .....	2003
emr-5.11.4 .....	2019
emr-5.11.3 .....	2034
emr-5.11.2 .....	2049
emr-5.11.1 .....	2064
emr-5.11.0 .....	2078
emr-5.10.1 .....	2094
emr-5.10.0 .....	2109
emr-5.9.1 .....	2125
emr-5.9.0 .....	2139
emr-5.8.3 .....	2155
emr-5.8.2 .....	2170
emr-5.8.1 .....	2184
emr-5.8.0 .....	2198
emr-5.7.1 .....	2214
emr-5.7.0 .....	2229
emr-5.6.1 .....	2243
emr-5.6.0 .....	2258
emr-5.5.4 .....	2272
emr-5.5.3 .....	2286
emr-5.5.2 .....	2301
emr-5.5.1 .....	2315
emr-5.5.0 .....	2329

emr-5.4.1 .....	2344
emr-5.4.0 .....	2358
emr-5.3.2 .....	2373
emr-5.3.1 .....	2387
emr-5.3.0 .....	2401
emr-5.2.3 .....	2416
emr-5.2.2 .....	2430
emr-5.2.1 .....	2444
emr-5.2.0 .....	2458
emr-5.1.1 .....	2473
emr-5.1.0 .....	2487
emr-5.0.3 .....	2501
emr-5.0.2 .....	2515
emr-5.0.1 .....	2528
emr-5.0.0 .....	2541
Amazon EMR 4.x リリースバージョン .....	2556
Amazon EMR 4.x リリースのアプリケーションバージョン .....	2557
リリースバージョンの違い .....	2558
emr-4.9.6 .....	2605
emr-4.9.5 .....	2619
emr-4.9.4 .....	2633
emr-4.9.3 .....	2646
emr-4.9.2 .....	2660
emr-4.9.1 .....	2674
emr-4.8.5 .....	2688
emr-4.8.4 .....	2702
emr-4.8.3 .....	2715
emr-4.8.2 .....	2729
emr-4.8.1 .....	2743
emr-4.8.0 .....	2757
emr-4.7.4 .....	2771
emr-4.7.3 .....	2784
emr-4.7.2 .....	2797
emr-4.7.1 .....	2810
emr-4.7.0 .....	2823
emr-4.6.1 .....	2837

emr-4.6.0 .....	2849
emr-4.5.0 .....	2862
emr-4.4.0 .....	2873
emr-4.3.0 .....	2885
emr-4.2.0 .....	2896
emr-4.1.0 .....	2906
emr-4.0.0 .....	2915
2.x および 3.x AMI バージョン .....	2922
クラスターの作成 .....	2923
アプリケーションのインストール .....	2925
設定のカスタマイズ .....	2925
[Hive] .....	2933
HBase .....	2944
Pig .....	2956
Spark .....	2962
S3DistCp .....	2966
最新情報 .....	2969
Amazon EMR 7.1.0 .....	2969
Amazon EMR 6.15.0 .....	2973
Amazon EMR 5.36.2 .....	2991
SigV4 の互換性 .....	2993
CVE-2021-44228 を軽減するためのアプローチ .....	2994
Log4j CVE-2021-44228 および CVE-2021-45046 用 Amazon EMR ブートストラップアク ションソリューション .....	2995
よくある質問 .....	3002
アーカイブ .....	3005
リリース 6.14.0 .....	3005
リリース 6.13.0 .....	3024
リリース 6.12.0 .....	3050
リリース 6.11.1 .....	3085
リリース 6.11.0 .....	3114
リリース 6.10.0 .....	3125
リリース 6.9.0 .....	3140
リリース 6.8.0 .....	3159
リリース 6.7.0 .....	3178
リリース 6.6.0 .....	3212

リリース 5.35.0 .....	3248
リリース 5.34.0 .....	3252
リリース 6.5.0 .....	3254
リリース 6.4.0 .....	3257
リリース 5.32.0 .....	3265
リリース 6.2.0 .....	3270
リリース 5.31.0 .....	3278
リリース 6.1.0 .....	3283
リリース 6.0.0 .....	3289
リリース 5.30.1 .....	3294
リリース 5.30.0 .....	3299
リリース 5.29.0 .....	3304
リリース 5.28.1 .....	3305
リリース 5.28.0 .....	3307
リリース 5.27.0 .....	3309
リリース 5.26.0 .....	3311
リリース 5.25.0 .....	3314
リリース 5.24.1 .....	3317
リリース 5.24.0 .....	3318
リリース 5.23.0 .....	3320
リリース 5.22.0 .....	3322
リリース 5.21.1 .....	3325
リリース 5.21.0 .....	3326
リリース 5.20.0 .....	3329
リリース 5.19.0 .....	3333
リリース 5.18.0 .....	3334
リリース 5.17.1 .....	3335
リリース 5.17.0 .....	3336
リリース 5.16.0 .....	3337
リリース 5.15.0 .....	3338
リリース 5.14.1 .....	3339
リリース 5.14.0 .....	3340
リリース 5.13.0 .....	3342
リリース 5.12.2 .....	3342
リリース 5.12.1 .....	3343
リリース 5.12.0 .....	3343

リリース 5.11.3 .....	3345
リリース 5.11.2 .....	3345
リリース 5.11.1 .....	3345
リリース 5.11.0 .....	3346
リリース 5.10.0 .....	3347
リリース 5.9.0 .....	3349
リリース 5.8.2 .....	3351
リリース 5.8.1 .....	3351
リリース 5.8.0 .....	3352
リリース 5.7.0 .....	3354
リリース 5.6.0 .....	3355
リリース 5.5.3 .....	3355
リリース 5.5.2 .....	3356
リリース 5.5.1 .....	3356
リリース 5.5.0 .....	3356
リリース 5.4.0 .....	3358
リリース 5.3.1 .....	3359
リリース 5.3.0 .....	3359
リリース 5.2.2 .....	3360
リリース 5.2.1 .....	3360
リリース 5.2.0 .....	3361
リリース 5.1.0 .....	3362
リリース 5.0.3 .....	3363
リリース 5.0.0 .....	3363
リリース 4.9.5 .....	3365
リリース 4.9.4 .....	3365
リリース 4.9.3 .....	3365
リリース 4.9.2 .....	3366
リリース 4.9.1 .....	3366
リリース 4.8.4 .....	3366
リリース 4.8.3 .....	3367
リリース 4.8.2 .....	3368
リリース 4.8.0 .....	3368
リリース 4.7.2 .....	3369
リリース 4.7.1 .....	3370
リリース 4.7.0 .....	3370

リリース 4.6.0 .....	3372
リリース 4.5.0 .....	3373
リリース 4.4.0 .....	3374
リリース 4.3.0 .....	3376
リリース 4.2.0 .....	3378
アプリケーションの設定 .....	3379
クラスターの作成時のアプリケーションの設定 .....	3381
クラスター作成時にコンソールで設定を指定する .....	3382
クラスターの作成 AWS CLI 時に を使用して設定を指定する .....	3382
クラスター作成時に Java SDK を使用して設定を指定する .....	3383
実行中のクラスター内のインスタンスグループの再設定 .....	3383
インスタンスグループの再設定時の考慮事項 .....	3384
コンソールでのインスタンスグループの再設定 .....	3387
CLI を使用したインスタンスグループの再設定 .....	3388
Java SDK を使用したインスタンスグループの再設定 .....	3393
トラブルシューティング .....	3394
機密設定データを AWS Secrets Manager に保存する .....	3397
シークレットを作成する .....	3397
Amazon EMR にアクセス権を付与してシークレットを取得する .....	3397
シークレットを設定分類で使用する .....	3398
シークレット値を更新する .....	3399
特定の Java 仮想マシンを使用するようにアプリケーションを設定 .....	3399
考慮事項 .....	3399
JVM をオーバーライドする .....	3402
サービスポート .....	3405
アプリケーションユーザー .....	3406
アーティファクトリポジトリを使用して依存関係を確認する .....	3408
EMR ファイルシステム (EMRFS) .....	3411
整合性のあるビュー .....	3413
整合性のあるビューを有効にする .....	3417
EMRFS の整合性のあるビューが Amazon S3 内のオブジェクトを追跡する方法を理解する .....	3419
再試行ロジック .....	3420
EMRFS の整合性のあるビューのメタデータ .....	3422
CloudWatch および Amazon SQS の整合性通知を設定する .....	3425
整合性のあるビューを設定する .....	3427

EMRFS CLI コマンドリファレンス .....	3431
Amazon S3 内の EMRFS データへのアクセスを許可する .....	3442
Amazon S3 内の EMRFS データに対してカスタムの認証情報プロバイダーを作成する ...	3442
デフォルト AWS Security Token Service エンドポイントの管理 .....	3444
EMRFS プロパティを使用して Amazon S3 の暗号化を指定する .....	3445
EMRFS 暗号化 AWS KMS keys に を使用する .....	3446
Amazon S3 のサーバー側の暗号化 .....	3447
Amazon S3 クライアント側の暗号化 .....	3450
Amazon CloudWatch エージェント .....	3458
クラスターを作成する .....	3460
必要なアクセス許可 .....	3460
必要な エンドポイント .....	3461
クラスターを作成する .....	3461
デフォルトメトリクス .....	3462
構成 .....	3465
Amazon EMR 7.1.0 .....	3465
Amazon EMR 7.0.0 .....	3499
考慮事項 .....	3504
履歴 .....	3505
Delta Lake .....	3506
序章 .....	3506
Delta Lake クラスターを使用する .....	3507
Flink で Delta Lake を使用する .....	3507
Trino で Delta Lake を使用する .....	3512
Spark で Delta Lake を使用する .....	3513
Spark および Glue と共に Delta Lake を使用する .....	3518
考慮事項 .....	3519
履歴 .....	3520
Flink .....	3522
Flink を使用してクラスターを作成する .....	3525
Flink の設定 .....	3526
Hive Metastore と Glue Catalog .....	3526
設定ファイル .....	3528
複数のプライマリノード .....	3529
メモリプロセスのサイズ .....	3530
ログ出カファイルのサイズ .....	3531

Java 11 .....	3532
Flink のジョブ .....	3537
長期実行のクラスターのステップとして、Flink YARN アプリケーションを起動します ....	3537
長期実行クラスター上の既存の Flink アプリケーションに作業を送信する .....	3539
一時的な Flink ジョブを送信する .....	3540
Flink Scala シェル .....	3542
Flink UI .....	3543
Flink オートスケーラー .....	3544
概要 .....	3544
考慮事項 .....	3545
オートスケーラーを有効にする .....	3545
Configurations .....	3547
再起動時間を最適化する .....	3551
タスクローカルリカバリ .....	3552
インクリメンタルチェックポイント .....	3553
きめ細かなリカバリ .....	3554
複合再起動 .....	3555
Zeppelin による Flink の操作 .....	3555
序章 .....	3555
前提条件 .....	3556
EMR クラスターで Zeppelin と Flink の連携を設定する .....	3556
Zeppelin と Flink が連携する EMR クラスターで Flink ジョブを実行する .....	3558
Flink リリース履歴 .....	3563
Ganglia .....	3601
Ganglia を使用したクラスターの作成 .....	3603
Ganglia メトリクスを表示する .....	3604
Ganglia での Hadoop と Spark のメトリクス .....	3605
Ganglia リリース履歴 .....	3606
Hadoop .....	3652
Hadoop の設定 .....	3654
タスクの設定 .....	3654
Hadoop デーモン構成設定 .....	4015
HDFS 構成 .....	4286
Amazon EMR における HDFS での透過的暗号化 .....	4287
HDFS 透過的暗号化を設定する .....	4288
HDFS 透過的暗号化の考慮事項 .....	4291



Hadoop キー管理サーバー .....	4291
複数のマスターノードを持つ EMR クラスターでの HDFS 透過的暗号化 .....	4295
Hadoop アプリケーションの作成または実行 .....	4296
Amazon EMR を使用してバイナリを構築する .....	4297
ストリーミングでのデータ処理 .....	4299
カスタム JAR でのデータの処理 .....	4305
YARN コンテナの不均一なメモリアクセス対応を有効にする .....	4309
Hadoop のバージョン履歴 .....	4311
Hadoop リリースノート (バージョン別) .....	4372
HBase .....	4376
HBase を使用するクラスターの作成 .....	4380
コンソールを使用して HBase を含むクラスターを作成する .....	4380
を使用して HBase でクラスターを作成する AWS CLI .....	4380
Amazon S3 の HBase (Amazon S3 ストレージモード) .....	4381
Amazon S3 での HBase の有効化 .....	4383
リードレプリカクラスターの使用 .....	4383
永続的 HFile トラッキング .....	4385
運用上の考慮事項 .....	4387
Amazon EMR のログ先行書き込み (WAL) .....	4391
WAL ワークスペース .....	4392
必要なアクセス許可 .....	4393
WAL の有効化 .....	4394
WAL からの復元 .....	4396
セキュリティ設定 .....	4397
の使用 AWS PrivateLink .....	4398
WAL の料金とメトリクス .....	4400
WAL ワークスペースのタグ付け .....	4401
考慮事項と可用性 .....	4402
EMRWAL CLI リファレンス .....	4404
HBase シェルを使用する .....	4407
テーブルを作成する .....	4408
値を入力する .....	4408
値を取得する .....	4408
テーブルを削除する .....	4408
Hive で HBase テーブルにアクセスする .....	4408
HBase のスナップショットを使用する .....	4410

テーブルを使用してスナップショットを作成する .....	4410
スナップショットを削除する .....	4411
スナップショット情報を表示する .....	4411
Amazon S3 にスナップショットをエクスポートする .....	4411
Amazon S3 からスナップショットをインポートする .....	4412
HBase シェル内のスナップショットからテーブルを復元する .....	4413
HBase を設定する .....	4414
YARN でメモリ割り当てを変更する .....	4415
HBase のポート番号 .....	4415
最適化のための HBase サイト設定 .....	4416
HBase ユーザーインターフェイスを表示する .....	4418
HBase ログファイルを表示する .....	4420
Ganglia で HBase を監視する .....	4420
HBase の以前のバージョンからの移行 .....	4422
HBase リリース履歴 .....	4422
HCatalog .....	4494
HCatalog を使用したクラスターの作成 .....	4495
HCatalog の使用 .....	4496
HCatalog HStorer の使用時に直接書き込みを無効にする .....	4497
HCat CLI を使用してテーブルを作成しそのデータを Pig で使用する .....	4497
Spark SQL を使用してテーブルにアクセスする .....	4498
例: HCatalog テーブルを作成して Pig を使用して書き込む .....	4500
HCatalog リリース履歴 .....	4501
[Hive] .....	4569
Amazon EMR 上の Hive の相違点と考慮事項 .....	4572
Amazon EMR 上の Apache Hive と Apache Hive の相違点 .....	4572
Amazon EMR リリースバージョン 4.x と 5.x の Hive の相違点 .....	4573
Amazon EMR での Hive の追加機能 .....	4574
Hive の外部メタストアの設定 .....	4579
AWS Glue データカタログを Hive のメタストアとして使用する .....	4580
外部の MySQL データベースまたは Amazon Aurora の使用 .....	4587
Hive JDBC ドライバーの使用 .....	4590
Hive のパフォーマンスの向上 .....	4592
Hive EMRFS S3 向けに最適化されたコミッターを有効にする .....	4592
S3 Select を使用する .....	4594
MSCK の最適化 .....	4596

Hive LLAP を使用する .....	4597
LLAP を有効にする .....	4597
クラスターで LLAP を起動する .....	4599
LLAP のステータスを確認する .....	4599
LLAP を起動または停止する .....	4600
LLAP デーモンの数を変更する .....	4600
Hive での暗号化 .....	4600
Hive の Parquet モジュール暗号化 .....	4600
HS2 での送信中の暗号化 .....	4604
Hive リリース履歴 .....	4605
Hive リリースノート (バージョン別) .....	4675
Hudi .....	4755
Hudi の仕組み .....	4757
データセットストレージタイプについて: 書き込み時にコピーおよび読み取り時にマージ	4757
メタストアへの Hudi データセットの登録 .....	4759
考慮事項と制約事項 .....	4759
Hudi がインストールされたクラスターを作成する .....	4761
Hudi データセットを操作する .....	4762
Hudi の Spark セッションを初期化する .....	4765
Hudi データセットに書き込む .....	4766
データのアップサート .....	4771
レコードを削除する .....	4772
Hudi データセットから読み込む .....	4774
Hudi CLI を使用する .....	4776
Hudi リリース履歴 .....	4776
Hue .....	4780
Hue バージョン .....	4780
Amazon EMR で Hue のサポートされている機能とサポート外の機能 .....	4782
考慮事項 .....	4783
大きな Hue メタデータテーブルでのパフォーマンス .....	4784
Hue バージョン間の非互換性 .....	4785
Hue Web ユーザーインターフェイスへの接続 .....	4785
Amazon RDS でリモートデータベースと Hue を使用する .....	4786
トラブルシューティング .....	4788
Hue の詳細設定 .....	4788
LDAP ユーザーに関して Hue を設定する .....	4789

Hue リリース履歴 .....	4792
Iceberg .....	4854
Iceberg の仕組み .....	4855
Iceberg でクラスターを使用する .....	4857
Spark で Iceberg クラスターを使用する .....	4857
Trino での Iceberg クラスターの使用 .....	4862
Flink での Iceberg クラスターの使用 .....	4864
Hive での Iceberg クラスターの使用 .....	4869
考慮事項と制約事項 .....	4872
Spark で Iceberg を使用するための考慮事項 .....	4872
Trino で Iceberg を使用するための考慮事項 .....	4872
Flink で Iceberg を使用するための考慮事項 .....	4873
Hive で Iceberg を使用するための考慮事項 .....	4873
Iceberg リリース履歴 .....	4873
バージョン別の Iceberg リリースノート .....	4875
Jupyter Notebook .....	4877
EMR Studio .....	4877
EMR ノートブック .....	4877
JupyterHub .....	4877
でクラスターを作成する JupyterHub .....	4882
Amazon EMR で を使用する場合 JupyterHub の考慮事項 .....	4884
の設定 JupyterHub .....	4885
Amazon S3 でノートブックの永続性を設定するには .....	4886
マスターノードとノートブックサーバーに接続する .....	4887
JupyterHub 設定と管理 .....	4888
Jupyter Notebook のユーザーと管理者を追加する .....	4889
追加のカーネルとライブラリをインストールする .....	4901
JupyterHub リリース履歴 .....	4905
Livy .....	4941
HTTPS を有効にする .....	4943
Livy リリース履歴 .....	4944
MXNet .....	4988
MXNet リリース履歴 .....	4990
Oozie .....	5018
Amazon RDS でリモートデータベースと Oozie を使用する .....	5020
Oozie 用の Java バージョンの設定 .....	5023

Oozie リリース履歴 .....	5023
バージョン別の Oozie リリースノート .....	5080
フェニックス .....	5082
Phoenix を使用したクラスターの作成 .....	5085
Phoenix 設定のカスタマイズ .....	5086
Phoenix クライアント .....	5087
Phoenix リリース履歴 .....	5090
Pig .....	5164
Pig の作業の送信 .....	5166
Amazon EMR コンソールを使用した Pig の作業の送信 .....	5167
を使用して Pig 作業を送信する AWS CLI .....	5168
Pig からユーザー定義関数を呼び出す .....	5169
Pig から JAR ファイルを呼び出す .....	5170
Pig から Python/Jython スクリプトを呼び出す .....	5170
Pig リリース履歴 .....	5171
Presto と Trino .....	5237
AWS Glue データカタログでの Presto の使用 .....	5240
AWS Glue Data Catalog をメタストアとして指定する .....	5241
IAM アクセス許可 .....	4584
AWS Glue Data Catalog を使用する場合の考慮事項 .....	5246
S3 Select Pushdown の使用 .....	5247
S3 Select Pushdown が使用するアプリケーションに適しているかどうかを確認する .....	5247
考慮事項と制約事項 .....	5248
PrestoDB または Trino で S3 Select Pushdown を有効にする .....	5248
データベースコネクタの追加 .....	5249
SSL/TLS と LDAPS の使用 .....	5250
LDAP 認証の使用 .....	5251
Presto 厳格モードを有効にする .....	5259
考慮事項 .....	5261
Presto でのスポットインスタンス損失の処理 .....	5261
耐障害性実行 .....	5263
構成 .....	5263
交換マネージャー .....	5264
考慮事項と制約事項 .....	5265
グレースフルな廃止による Presto Auto Scaling の使用 .....	5266
Amazon EMR での Presto に関する考慮事項 .....	5267

Presto コマンドライン実行可能ファイル .....	5267
Presto デプロイプロパティは設定可能ではない .....	5267
PrestoDB と Trino のインストール .....	5269
EMRFS と PrestoS3FileSystem の設定 .....	5269
エンドユーザー偽装のデフォルト設定 .....	5270
Presto ウェブインターフェイス用のデフォルトポート .....	5270
一部のリリースでの Hive バケットの実行に関する問題 .....	5270
Presto リリース履歴 .....	5271
Trino (PrestoSQL) リリースノート (バージョン別) .....	5335
Spark .....	5338
Spark クラスターを作成する .....	5342
Amazon EMR 6.x を使用して Docker で Spark アプリケーションを実行する .....	5345
考慮事項 .....	5345
Docker イメージの作成 .....	5346
Amazon ECR の Docker イメージを使用する .....	5347
AWS Glue データカタログを Spark SQL のメタストアとして使用する .....	5352
AWS Glue Data Catalog をメタストアとして指定する .....	5353
IAM アクセス許可 .....	4584
考慮事項 .....	4585
Spark の設定 .....	5358
Spark のデフォルト .....	5359
ガベージコレクションの設定 .....	5361
maximizeResourceAllocation .....	5361
廃止動作 .....	5364
Spark ThriftServer 環境変数 .....	5367
Spark のデフォルト設定の変更 .....	5367
Apache Log4j 1.x から Log4j 2.x への移行 .....	5370
Spark パフォーマンスの最適化 .....	5370
アダプティブクエリ実行 .....	5370
ダイナミックパーティションプルーニング .....	5373
スカラサブクエリの平坦化 .....	5375
INTERSECT 前に DISTINCT .....	5376
ブルームフィルター結合 .....	5377
結合順序の最適化 .....	5377
結果フラグメントキャッシュ .....	5378
Spark 結果フラグメントキャッシュを有効にする .....	5379

考慮事項 .....	5379
RAPIDS アクセラレーター .....	5381
インスタンスタイプの選択 .....	5381
アプリ設定をセットアップする .....	5382
ブートストラップアクションを追加する .....	5389
クラスターを起動する .....	5390
Spark シェルにアクセスする .....	5390
機械学習に Amazon SageMaker Spark を使用する .....	5392
Spark アプリケーションを作成する .....	5393
Scala .....	5393
Java .....	5394
Python .....	5395
S3 で Spark のパフォーマンスを向上させる .....	5396
S3 Select を使用する .....	5397
EMRFS S3 用に最適化されたコミッター .....	5401
EMRFS S3 向けに最適化されたコミットプロトコルを使用する .....	5408
S3 リクエストを再試行する .....	5414
Spark ステップを追加する .....	5417
Spark のデフォルト設定を上書きする .....	5421
Spark アプリケーション履歴を表示する .....	5421
Spark ウェブ UI にアクセスする .....	5421
Amazon Kinesis Data Streams での Spark の使用 .....	5422
Amazon Redshift での Spark の使用 .....	5422
Spark アプリケーションの起動 .....	5423
Amazon Redshift の認証 .....	5424
Amazon Redshift に対する読み書き .....	5426
考慮事項 .....	5428
Spark リリース履歴 .....	5429
Sqoop .....	5492
Sqoop バージョン .....	5492
Amazon EMR での Sqoop に関する考慮事項 .....	5494
HCatalog の統合での Sqoop の使用 .....	5494
Sqoop JDBC およびデータベースサポート .....	5494
パスワードの保護 .....	5496
Sqoop リリース履歴 .....	5497
TensorFlow .....	5550

TensorFlow Amazon EC2 インスタンスタイプ別の ビルド .....	5552
セキュリティ .....	5552
の使用 TensorBoard .....	5552
TensorFlow リリース履歴 .....	5553
Tez .....	5576
Tez を使用したクラスターの作成 .....	5578
Tez の設定 .....	5579
設定例 .....	5579
Tez 非同期分割オープニング .....	5580
Tez ウェブ UI .....	5582
タイムラインサーバー .....	5583
Tez リリース履歴 .....	5583
Tez リリースノート (バージョン別) .....	5627
Zeppelin .....	5634
Amazon EMR で Zeppelin を使用するときの考慮事項 .....	5636
Zeppelin リリース履歴 .....	5637
ZooKeeper .....	5695
ZooKeeper リリース履歴 .....	5697
コネクタおよびユーティリティ .....	5736
DynamoDB 内テーブルのエクスポート、クエリ、結合 .....	5736
Hive コマンドを実行するように Hive テーブルをセットアップする .....	5738
データのエクスポート、インポート、クエリを行う Hive コマンドの使用例 .....	5747
パフォーマンスの最適化 .....	5756
Kinesis .....	5760
Amazon EMR と Amazon Kinesis の統合で何ができますか? .....	5760
Amazon Kinesis Streams のチェックポイントの分析 .....	5760
パフォーマンスに関する考慮事項 .....	5762
Amazon EMR で Amazon Kinesis 分析をスケジュールする .....	5762
Amazon EMR 7.0 用スパークキネシスコネクタの SDK 2.x への移行 .....	5762
S3DistCp (s3-dist-cp) .....	5771
S3DistCp オプション .....	5772
S3DistCp をクラスター内のステップとして追加する .....	5780
S3DistCp ジョブが失敗した後のクリーンアップ .....	5782
クラスターでのコマンドとスクリプトの実行 .....	5784
スクリプトまたはコマンドを実行するためのカスタム JAR ステップの送信 .....	5784
command-runner.jar のその他の使用方法 .....	5786



---

AWS 用語集 .....	5788
.....	5789

# Amazon EMR リリースについて

Amazon EMR リリースは、ビッグデータエコシステムからの一連のオープンソースのアプリケーションです。各リリースは異なるビッグデータアプリケーション、コンポーネント、および機能で構成され、クラスターを作成するときに Amazon EMR でインストールして設定することを選択します。アプリケーションは、Hadoop エコシステムに関連付けられたオープンソースプロジェクトである [Apache BigTop](#) に基づくシステムを使用してパッケージ化されます。このガイドでは、Amazon EMR リリースに含まれるアプリケーションの情報を提供します。

Amazon EMR の開始方法および使用方法の詳細については、[「Amazon EMR 管理ガイド」](#)を参照してください。

クラスターを起動するときに、複数の Amazon EMR リリースから選択できます。これにより、互換性の要件に適合するアプリケーションバージョンをテストして使用できます。リリース番号はリリースラベルで指定します。リリースラベルの形式は `emr-x.x.x` です。例えば `emr-7.1.0` です。

Amazon EMR 5.18.0 以降では、Amazon EMR アーティファクトリポジトリを使用し、特定の Amazon EMR リリースで利用可能なライブラリおよび依存関係の正確なバージョンに対してジョブコードをビルドできます。詳細については、[「Amazon EMR アーティファクトリポジトリを使用して依存関係を確認する」](#)を参照してください。

新しい Amazon EMR リリースが利用可能になったときに更新を取得するには、[Amazon EMR リリースノートの RSS フィード](#)にサブスクライブします。

最新リリースの詳細。Amazon EMR 7.x、6 x および 5.x シリーズのアプリケーションバージョン、リリースノート、コンポーネント、設定分類が含まれます。

- [Amazon EMR リリース 7.1.0](#)
- [Amazon EMR リリース 6.15.0](#)
- [Amazon EMR リリース 5.36.2](#)

## Note

新しい Amazon EMR リリースは、特定のリージョンで最初に利用可能になり、その後数日で、他のリージョンでも利用できるようになります。この期間中、お客様のリージョンで最新のリリースバージョンが利用できない場合があります。

最新の Amazon EMR リリースのリリースノートおよびすべてのリリースの履歴:

- [最新情報](#)
- [Amazon EMR リリースノートのアーカイブ](#)

各 Amazon EMR リリースの包括的なアプリケーションバージョン履歴:

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

各 Amazon EMR リリースの詳細およびリリースシリーズ間の違い (該当する場合)。

- [Amazon EMR 7.x リリースバージョン](#)
- [Amazon EMR 6.x リリースバージョン](#)
- [Amazon EMR 5.x リリースバージョン](#)
- [Amazon EMR 4.x リリースバージョン](#)
- [Amazon EMR 2.x および 3.x AMI バージョン](#)

## Amazon EMR 7.x リリースバージョン

このセクションでは、各 Amazon EMR 7.x リリースバージョンのアプリケーションバージョン、リリースノート、コンポーネントバージョン、設定分類を示しています。

クラスターを起動するときに、複数の Amazon EMR リリースから選択できます。これにより、互換性の要件に適合するアプリケーションバージョンをテストして使用できます。リリース番号はリリースラベルで指定します。リリースラベルの形式は `emr-x.x.x` です。例えば `emr-7.1.0` です。

新しい Amazon EMR リリースは、特定のリージョンで最初に利用可能になり、その後数日で、他のリージョンでも利用できるようになります。この期間中、お客様のリージョンで最新のリリースバージョンが利用できない場合があります。

すべての Amazon EMR 7.x リリースのアプリケーションバージョンの包括的な表については、「[Amazon EMR 7.x リリースのアプリケーションバージョン](#)」を参照してください。

トピック

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR リリース 7.1.0](#)
- [Amazon EMR リリース 7.0.0](#)

## Amazon EMR 7.x リリースのアプリケーションバージョン

各 Amazon EMR 7.x リリースで利用可能なアプリケーションバージョンを示した包括的な表については、ブラウザで「[Amazon EMR 7.x リリースのアプリケーションバージョン](#)」を開いて確認してください。

### Amazon EMR リリース 7.1.0

#### 7.1.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[AmazonCloudWatchAgent](#)、[Delta](#)、[Flink](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[Jupyter](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.23.18、 1.12.656	2.20.160-amzn-0、 1.12.569
Python	3.9	3.9
Scala	2.12.17	2.12.17

	emr-7.1.0	emr-7.0.0
AmazonCloudWatchAgent	1.300032.2	1.300031.1
Delta	3.0.0	3.0.0
Flink	1.18.1	1.18.0
Ganglia	-	-
HBase	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6
Hive	3.1.3	3.1.3
Hudi	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0
Iceberg	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0
Livy	0.8.0	0.7.1
MXNet	1.9.1	1.9.1
Mahout	-	-
Oozie	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3
Pig	0.17.0	0.17.0
Presto	0.284	0.283

	emr-7.1.0	emr-7.0.0
Spark	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0
Tez	0.10.2	0.10.2
Trino (PrestoSQL)	435	426
Zeppelin	0.10.1	0.10.1
ZooKeeper	3.9.1	3.5.10

## 7.1.0 リリースノート

次のリリースノートには、Amazon EMR リリース 7.1.0 に関する情報が含まれています。7.1.0 からの変更が含まれています。

### 新機能

- アプリケーションのアップグレード - Amazon EMR 7.1.0 application upgrades include Livy 0.8.0, Trino 435, and ZooKeeper 3.9.1。
- [異常のあるノードの置換](#) — Amazon EMR 7.1.0 以降では、異常のあるノードの置換がデフォルトで有効になっているため、Amazon EMR は異常のあるノードを適切に置き換えます。Amazon EMR リリース 7.0.0 以前の既存のワークフローに影響を与えないように、クラスターで終了保護を有効にした場合、異常なノード交換は無効になります。
- [CloudWatch エージェント](#) – Amazon EMR 設定 API を使用して、追加のシステムメトリクスを使用するように CloudWatch エージェントを設定し、アプリケーションメトリクスを追加し、メトリクスの送信先を変更します。

### 既知の問題

- Python 3.11 は EMR Studio ではサポートされていません。

## 変更点、機能強化、解決した問題

- Amazon EMR 7.1.0 はデフォルトで Python 3.9 をサポートしていますが、Amazon EMR 7.1.0 の Livy 0.8.0 と Spark は Python 3.11 をサポートしています。
- このリリースでは、Python バージョン 3.11 PySpark でを使用する場合に、各行を 1 つずつ実行する必要があるという問題が修正されています。
- Zeppelin のアップグレード – Amazon EMR 7.1.0 には、Zeppelin の AWS SDK for Java v2 へのアップグレードが含まれています。このアップグレードにより、Zeppelin S3 ノートブックはカスタム暗号化マテリアルプロバイダーを受け入れることができます。AWS SDK for Java v2 は EncryptionMaterialsProvider インターフェイスを削除します。Amazon EMR 7.1.0 にアップグレードする場合、カスタム暗号化を使用する場合は、[キーリングインターフェイス](#)を実装する必要があります。Keyring インターフェイスの実装方法の例については、[KmsKeyring「.java」](#)を参照してください。
- Amazon EMR リリース 7.1.0 にアップグレードする場合は、ローカルディスク暗号化用のカスタムキープロバイダーを変更して、AES/GCM/ の AES アルゴリズムを使用してキーを生成します NoPadding。アルゴリズムを更新しない場合、クラスターの作成が失敗してエラーが発生する可能性があります Local disk encryption failed on master instance (i-123456789) due to internal error.。カスタムキープロバイダーの作成の詳細については、[「カスタムキープロバイダーの作成」](#)を参照してください。
- Amazon EMR 7.1.0 は、開いているファイルハンドルを持つファイルのログ切り捨てロジックを改善することで、ディスク容量が少ない条件下でのノードの回復性を向上させます。
- このリリースでは、ノードの再起動時に Amazon EMR デーモンの読み取りおよび書き込みファイルを使用して、エンコーディングとデコードのロジックを強化し、データ破損やノード障害のリスクを最小限に抑えます。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、[「Amazon EMR にデフォルトの Amazon Linux AMI を使用する」](#)を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2023年3月2024日219.0	6.1.77-99.164.amzn2023	2024年5月8日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、カナダ西部 (カルガリー)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

### 7.1.0 のデフォルト Java バージョン

Amazon EMR リリース 7.1 以降には、Apache Livy を除く Corretto 17 (OpenJDK) をサポートするアプリケーション用に、デフォルトで Amazon Corretto 17 (OpenJDK 上に構築) が付属しています。

次の表は、Amazon EMR 7.1.0 のアプリケーションのデフォルトの Java バージョンを示しています。クラスターのデフォルト JVM を変更する場合は、クラスター上で実行される各アプリケーションに対して「[特定の Java 仮想マシンを使用するようにアプリケーションを設定](#)」の手順を実行してください。1 つのクラスターに使用できる Java ランタイムバージョンは 1 つだけです。Amazon EMR は、同じクラスター上の異なるランタイムバージョンで異なるノードやアプリケーションを実行することをサポートしていません。

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
差分	<b>17</b> 、11、8
Flink	17、 <b>11</b> 、8

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
Ganglia	8
HBase	11、8
HCatalog	17、11、8
Hadoop	17、11、8
[Hive]	17、11、8
Hudi	17、11、8
Iceberg	17、11、8
Livy	17、11、8
Oozie	17、11、8
フェニックス	8
PrestoDB	8
Spark	17、11、8
Spark RAPIDS	17、11、8
Sqoop	8
Tez	17、11、8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

## 7.1.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>adot-java-agent</code>	1.31.0	アプリケーションデーモンからメトリクスを収集する Java エージェント。
<code>delta</code>	3.0.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>emr-amazon-cloudwatch-agent</code>	1.300032.2-amzn-0	Amazon EC2 インスタンスから内部システムレベルのメトリクスとカスタムアプリケーションメトリクスを収集するアプリケーション。
<code>emr-ddb</code>	5.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.10.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.14.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.31.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.10.0	EMR S3Select コネクタ
emr-wal-cli	1.2.2	emrwal のリスト/削除に使用される CLI。
emrfs	2.62.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.18.1-amzn-0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.18.1-amzn-0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
hadoop-client	3.3.6-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.6-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.6-amzn-3	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	3.3.6-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.6-amzn-3	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.3.6-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.6-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.3.6-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.6-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.6-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.6-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.17-amzn-5	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。

コンポーネント	Version	説明
hbase-region-server	2.4.17-amzn-5	1つ以上のHBaseリージョンに対応するサービス。
hbase-client	2.4.17-amzn-5	HBase コマンドラインクライアント。
hbase-rest-server	2.4.17-amzn-5	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.17-amzn-5	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.17-amzn-5	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-10	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-10	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-10	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-10	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-10	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	3.1.3-amzn-10	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-10	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.14.1-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.14.1-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.14.1-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.14.1-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.4.3-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.5.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.8.0 インキュベーション	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.7.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.3	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.284-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。



コンポーネント	Version	説明
presto-worker	0.284-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.284-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	435-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	435-amzn-0	いくつかのクエリを実行するサービス。
trino-client	435-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.3.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.5.0-amzn-1	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	3.5.0-amzn-1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.5.0-amzn-1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.5.0-amzn-1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	23.10.0-amzn-1	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-8	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-8	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.9.1-amzn-0	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.9.1-amzn-0	ZooKeeper コマンドラインクライアント。

## 7.1.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-7.1.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
docker-conf	Docker 関連の設定を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。

分類	説明
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-env	HDFS 環境の値を変更します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive	Apache Hive 用の Amazon EMR-curated 設定。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。

分類	説明
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-java-home	Hadoop の KMS java ホームを変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。

分類	説明
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j2	Livy の log4j2.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。

分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-lakeformation	Presto の lakeformation.properties ファイルの値を変更します。



分類	説明
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
trino-log	Trino の log.properties ファイルの値を変更します。
trino-config	Trino の config.properties ファイルの値を変更します。
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。

分類	説明
trino-env	Trino の trino-env.sh ファイルの値を変更します。
trino-node	Trino の node.properties ファイルの値を変更します。
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。

分類	説明
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zeppelin-site	zeppelin-site.xml の構成設定を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-logback	ZooKeeperの logback.xml ファイルの値を変更します。
emr-metrics	このノードの emr メトリクス設定を変更します。

## 7.1.0 変更ログ

### 7.1.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2024-05-10	ドキュメント公開	Amazon EMR 7.1.0 リリースノートが初めて公開されました
2024-05-08	デプロイ完了	Amazon EMR 7.1.0 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2024-04-23	初回リリース	Amazon EMR 7.1.0 が最初の商用リージョンに初めてデプロイされました

## Amazon EMR リリース 7.0.0

### 7.0.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[AmazonCloudWatchAgent](#)、[Delta](#)、[Flink](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[Jupyter](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.23.18、 1.12.656	2.20.160-amzn-0、 1.12.569
Python	3.9	3.9
Scala	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300031.1
Delta	3.0.0	3.0.0
Flink	1.18.1	1.18.0
Ganglia	-	-
HBase	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6
Hive	3.1.3	3.1.3
Hudi	0.14.1-amzn-0	0.14.0-amzn-1

	emr-7.1.0	emr-7.0.0
Hue	4.11.0	4.11.0
Iceberg	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0
Livy	0.8.0	0.7.1
MXNet	1.9.1	1.9.1
Mahout	-	-
Oozie	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3
Pig	0.17.0	0.17.0
Presto	0.284	0.283
Spark	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0
Tez	0.10.2	0.10.2
Trino (PrestoSQL)	435	426
Zeppelin	0.10.1	0.10.1
ZooKeeper	3.9.1	3.5.10

## 7.0.0 リリースノート

次のリリースノートには、Amazon EMR リリース 7.0.0 に関する情報が含まれています。6.15.0 からの変更が含まれています。

### 新機能

- アプリケーションのアップグレード - Amazon EMR 7.0.0 application upgrades include Python 3.9, Spark 3.5, Flink 1.18, and Delta 3.0。また、このリリースでは、[Amazon CloudWatch エージェント](#) アプリケーションのサポートが追加され、Ganglia のサポートは削除されました。
- Amazon Corretto 17 - Amazon EMR リリース 7.0 以降には、Corretto 17 (JDK 17) をサポートするアプリケーションには、Apache Livy を除き、デフォルトで Amazon Corretto 17 (OpenJDK 上に構築) が付属しています。このリリースでサポートされている JDK バージョンの詳細については、「[7.0.0 のデフォルト Java バージョン](#)」を参照してください。
- Amazon Linux 2023 - 7.0 リリースでは、Amazon EMR クラスターはデフォルトで AL2023 で実行されるようになりました。これがデフォルトの AMI バージョンにどのように影響するかについては、「Amazon EMR 管理ガイド」の「[ソフトウェア更新に関する考慮事項](#)」を参照してください。また、AL2023 では Python 2.7 が削除されているため、Python を必要とするコンポーネントはすべて Python 3 で作成する必要があることにも注意してください。
- s3a を使用した Outposts での S3 - Amazon EMR は、s3a ファイルシステムを使用して Outposts バケットでの Amazon S3 をサポートするようになりました。Outposts 上の S3 の詳細については、Amazon Simple Storage Service ユーザーガイドの「[S3 on Outposts とは](#)」を参照してください。

### 既知の問題

- Python バージョン 3.10 以降 PySpark で を使用している場合、一度に複数の行を実行することはできません。各行は一度に 1 つずつ実行する必要があります。

### 変更点、機能強化、解決した問題

- クラスター内のすべてのインスタンスの状態を維持するために、Amazon EMR は Apache YARN、Apache HDFS、および Kerberos と統合されています。7.0 では、スケーリングやその他の操作により終了したインスタンスの状態を確実に削除できるように、これらの統合を強化しました。頻繁にスケーリングされ、その存続期間中に終了したインスタンスが何千も蓄積されるため、マネージドスケーリングが有効になっている長期間有効なクラスターでは特に重要です。



- このリリースでは、Kerberos 設定が改善され、AES ベースの暗号のみがサポートされるようになりました。AES ベースではない暗号を使用する Kerberos KDC では、Amazon EMR リリース 7.0.0 以降で実行される EMR クラスターではサポートされなくなりました。AES ベースの暗号は、クラスターに最強のセキュリティを提供します。
- AWS SDK 2.x 移行の一環として、Amazon EMR 7.0 には互換性のために Spark Kinesis コネクタの更新が含まれています。このアップデートは Apache Spark のコミュニティバージョンではご利用いただけません。7.0 より前の Amazon EMR リリースの Spark Kinesis コネクタを使用する場合、ワークロードを Amazon EMR 7.0 に移行する前に、アプリケーションコードを SDK 2.x で実行するように移行する必要があります。詳細については、「[Amazon EMR 7.0 用スパークキネシスコネクタの SDK 2.x への移行](#)」を参照してください。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2023年3月2024日0304.0	6.1.79-99.164.amzn2023	2024年3月12日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、ア

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			<p>アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、カナダ西部 (カルガリー)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)</p>

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2023 年 3 月 2024 日 219.0	6.1.77-99.164.amzn2023	2024 年 3 月 1 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			(バーレーン)、中東(アラブ首長国連邦)、カナダ(中部)、イスラエル(テルアビブ)、カナダ西部(カルガリー)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2023 年 3 月 2024 日 205.0	6.1.75-99.163.amzn2023	2024 年 2 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			(バーレーン)、中東(アラブ首長国連邦)、カナダ(中部)、イスラエル(テルアビブ)、カナダ西部(カルガリー)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2023 年 3 月 2024 日 122.0	6.1.72-96.166.amzn2023	2024 年 2 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			(バーレーン)、中東(アラブ首長国連邦)、カナダ(中部)、イスラエル(テルアビブ)、カナダ西部(カルガリー)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2023 年 3 月 2024 日 108.0	6.1.72-96.166.amzn2023	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			(バーレーン)、中東(アラブ首長国連邦)、カナダ(中部)、イスラエル(テルアビブ)、カナダ西部(カルガリー)、AWS GovCloud(米国西部)、AWS GovCloud(米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2023 年 3 月 2023 日 1211.4	6.1.66-91.160.amzn2023	2023 年 12 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

## 7.0.0 のデフォルト Java バージョン

Amazon EMR リリース 7.0 以降には、Corretto 17 (JDK 17) をサポートするアプリケーションには、Apache Livy を除き、デフォルトで Amazon Corretto 17 (OpenJDK 上に構築) が付属しています。

次の表は、Amazon EMR 7.0.0 のアプリケーションのデフォルト Java バージョンを示しています。クラスターのデフォルト JVM を変更する場合は、クラスター上で実行される各アプリケーションに対して「[特定の Java 仮想マシンを使用するようにアプリケーションを設定](#)」の手順を実行してください。1 つのクラスターに使用できる Java ランタイムバージョンは 1 つだけです。Amazon EMR は、同じクラスター上の異なるランタイムバージョンで異なるノードやアプリケーションを実行することをサポートしていません。

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
差分	<b>17</b> 、11、8
Flink	11、 <b>8</b>

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
Ganglia	8
HBase	11、8
HCatalog	17、11、8
Hadoop	17、11、8
[Hive]	17、11、8
Hudi	17、11、8
Iceberg	17、11、8
Livy	17、11、8
Oozie	17、11、8
フェニックス	8
PrestoDB	8
Spark	17、11、8
Spark RAPIDS	17、11、8
Sqoop	8
Tez	17、11、8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

## 7.0.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>adot-java-agent</code>	1.31.0	アプリケーションデーモンからメトリクスを収集する Java エージェント。
<code>delta</code>	3.0.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>delta-standalone-connectors</code>	0.6.0	Delta コネクタは、Flink、Hive、Presto などのエンジンと Delta Lake を統合するためのさまざまなランタイムを提供します。
<code>emr-amazon-cloudwatch-agent</code>	1.300031.1-amzn-0	Amazon EC2 インスタンスから内部システムレベルのメトリクスとカスタムアプリケーションメトリクスを収集するアプリケーション。

コンポーネント	Version	説明
emr-ddb	5.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.9.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.13.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.30.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.9.0	EMR S3Select コネクタ
emr-wal-cli	1.2.1	emrwal のリスト/削除に使用される CLI。
emrfs	2.61.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.18.0-amzn-0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.18.0-amzn-0	Apache Flink の EMR ノード上のリソースの管理 JobManager。

コンポーネント	Version	説明
hadoop-client	3.3.6-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.6-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.6-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.6-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.6-amzn-2	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httfs-server	3.3.6-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.6-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.3.6-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.6-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.6-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。



コンポーネント	Version	説明
hadoop-yarn-timeline-server	3.3.6-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.17-amzn-4	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.17-amzn-4	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.17-amzn-4	HBase コマンドラインクライアント。
hbase-rest-server	2.4.17-amzn-4	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.17-amzn-4	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.17-amzn-4	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-9	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-9	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-9	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。

コンポーネント	Version	説明
hive-client	3.1.3-amzn-9	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-9	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-9	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-9	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.14.0-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.14.0-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.14.0-amzn-1	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.14.0-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.4.2-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。

コンポーネント	Version	説明
jupyterhub	1.5.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.7.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.3	Spark-3 用 Apache Phoenix-Connectors

コンポーネント	Version	説明
phoenix-query-server	5.1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.283-amzn-1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.283-amzn-1	いくつかのクエリを実行するサービス。
presto-client	0.283-amzn-1	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	426-amzn-1	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	426-amzn-1	いくつかのクエリを実行するサービス。
trino-client	426-amzn-1	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.1.3	統計的コンピューティング用 R プロジェクト

コンポーネント	Version	説明
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.5.0-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.5.0-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.5.0-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.5.0-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	23.10.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-7	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-7	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.58	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

## 7.0.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-7.0.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
docker-conf	Docker 関連の設定を変更します。
emrfs-site	EMRFS の設定を変更します。

分類	説明
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。

分類	説明
hdfs-env	HDFS 環境の値を変更します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive	Apache Hive 用の Amazon EMR-curated 設定。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。



分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-java-home	Hadoop の KMS java ホームを変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。

分類	説明
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j2	Livy の log4j2.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-lakeformation	Presto の lakeformation.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。

分類	説明
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
trino-log	Trino の log.properties ファイルの値を変更します。
trino-config	Trino の config.properties ファイルの値を変更します。
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。
trino-env	Trino の trino-env.sh ファイルの値を変更します。
trino-node	Trino の node.properties ファイルの値を変更します。
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。

分類	説明
trino-connector-jmx	Trino の <code>jmx.properties</code> ファイルの値を変更します。
trino-connector-kafka	Trino の <code>kafka.properties</code> ファイルの値を変更します。
trino-connector-localfile	Trino の <code>localfile.properties</code> ファイルの値を変更します。
trino-connector-memory	Trino の <code>memory.properties</code> ファイルの値を変更します。
trino-connector-mongodb	Trino の <code>mongodb.properties</code> ファイルの値を変更します。
trino-connector-mysql	Trino の <code>mysql.properties</code> ファイルの値を変更します。
trino-connector-postgresql	Presto の <code>postgresql.properties</code> ファイルの値を変更します。
trino-connector-raptor	Trino の <code>raptor.properties</code> ファイルの値を変更します。
trino-connector-redis	Trino の <code>redis.properties</code> ファイルの値を変更します。
trino-connector-redshift	Trino の <code>redshift.properties</code> ファイルの値を変更します。
trino-connector-tpch	Trino の <code>tpch.properties</code> ファイルの値を変更します。
trino-connector-tpcds	Trino の <code>tpcds.properties</code> ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の <code>dbks-site.xml</code> ファイルの値を変更します。

分類	説明
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。

分類	説明
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zeppelin-site	zeppelin-site.xml の構成設定を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## 7.0.0 変更ログ

### 7.0.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-12-29	デプロイ完了	Amazon EMR 7.0.0 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-12-28	ドキュメント公開	Amazon EMR 7.0.0 リリースノートの初版が公開されました
2023-12-19	初回リリース	Amazon EMR 7.0.0 が初回商用リージョンに初めてデプロイされました



# Amazon EMR 6.x リリースバージョン

このセクションでは、各 Amazon EMR 6.x リリースバージョンで利用可能なアプリケーションバージョン、リリースノート、コンポーネントバージョン、および設定分類を示しています。

クラスターを起動するときに、複数の Amazon EMR リリースから選択できます。これにより、互換性の要件に適合するアプリケーションバージョンをテストして使用できます。リリース番号はリリースラベルで指定します。リリースラベルの形式は `emr-x.x.x` です。例えば `emr-7.1.0` です。

新しい Amazon EMR リリースは、特定のリージョンで最初に利用可能になり、その後数日で、他のリージョンでも利用できるようになります。この期間中、お客様のリージョンで最新のリリースバージョンが利用できない場合があります。

すべての Amazon EMR 6.x リリースのアプリケーションバージョンの包括的な表については、「[Amazon EMR 6.x リリースのアプリケーションバージョン](#)」を参照してください。

## トピック

- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR リリース 6.15.0](#)
- [Amazon EMR リリース 6.14.0](#)
- [Amazon EMR リリース 6.13.0](#)
- [Amazon EMR リリース 6.12.0](#)
- [Amazon EMR リリース 6.11.1](#)
- [Amazon EMR リリース 6.11.0](#)
- [Amazon EMR リリース 6.10.1](#)
- [Amazon EMR リリース 6.10.0](#)
- [Amazon EMR リリース 6.9.1](#)
- [Amazon EMR リリース 6.9.0](#)
- [Amazon EMR リリース 6.8.1](#)
- [Amazon EMR リリース 6.8.0](#)
- [Amazon EMR リリース 6.7.0](#)
- [Amazon EMR リリース 6.6.0](#)
- [Amazon EMR リリース 6.5.0](#)

- [Amazon EMR リリース 6.4.0](#)
- [Amazon EMR リリース 6.3.1](#)
- [Amazon EMR リリース 6.3.0](#)
- [Amazon EMR リリース 6.2.1](#)
- [Amazon EMR リリース 6.2.0](#)
- [Amazon EMR リリース 6.1.1](#)
- [Amazon EMR リリース 6.1.0](#)
- [Amazon EMR リリース 6.0.1](#)
- [Amazon EMR リリース 6.0.0](#)

## Amazon EMR 6.x リリースのアプリケーションバージョン

各 Amazon EMR 6.x リリースで利用可能なアプリケーションバージョンを示した包括的な表については、ブラウザで「[Amazon EMR 6.x リリースのアプリケーションバージョン](#)」を開いて確認してください。

### Amazon EMR リリース 6.15.0

#### 6.15.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
AWS SDK for Java	2.20.160-amzn-0、1.12.569	1.12.543	1.12.513	1.12.490
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.17	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.4.0
Flink	1.17.1	1.17.1	1.17.0	1.17.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.14.0-amzn-0	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.4.0-amzn-0	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.283	0.281	0.281	0.281
Spark	3.4.1	3.4.1	3.4.1	3.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	426	422	414	414
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

## 6.15.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.15.0 に関する情報が含まれています。6.14.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.15.0 変更ログ](#)」を参照してください。

### 新機能

- アプリケーションのアップグレード - Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426.

- [EC2 で実行される EMR クラスターの起動時間の短縮](#) - EC2 クラスターでの Amazon EMR の起動が最大 35% 速くなりました。この改善により、お客様の大多数はクラスターを 5 分以内に起動できます。
- [CodeWhisperer for EMR Studio](#) – Amazon CodeWhisperer EMR Studio で Amazon を使用して、コードを記述する際にリアルタイムのレコメンデーションを取得できるようになりました JupyterLab。CodeWhisperer は、コメントの入力、単一行のコードの完了、レコメンデーションの作成 line-by-line、および完全な形式の関数の生成を行うことができます。
- [Flink によるジョブの再起動時間の短縮](#) - Amazon EMR 6.15.0 以降では、Apache Flink にいくつかの新しいメカニズムが用意され、タスクの回復またはスケーリング操作中のジョブの再起動時間を短縮できます。これにより、実行グラフの回復と再開の速度を最適化してジョブの安定性を高めることができます。
- [オープンテーブル形式のテーブルレベルおよびきめ細かなアクセスコントロール](#) – Amazon EMR 6.15.0 以降では、AWS Glue Data Catalog のデータにアクセスする EC2 クラスターで Amazon EMR で Spark ジョブを実行すると、AWS Lake Formation を使用して Hudi、Iceberg、または Delta Lake ベースのテーブルにテーブル、行、列、およびセルレベルのアクセス許可を適用できます。
- Hadoop のアップグレード - Amazon EMR 6.15.0 には、Apache Hadoop のバージョン 3.3.6 へのアップグレードが含まれています。Hadoop 3.3.6 は、2023 年 6 月にアパッチによってリリースされた Amazon EMR 6.15 のデプロイ時の最新バージョンでした。Amazon EMR の以前のリリース (6.9.0 から 6.14.x) では Hadoop 3.3.3 が使用されていました。

このアップグレードには数百の改善と修正が含まれており、再構成可能なデータノードパラメータ、すべてのライブデータノードで一括再構成操作を開始する DFSAdmin オプション、シーク負荷の高いリーダーが複数の読み取り範囲を指定できるベクトル API などの機能が含まれています。Hadoop 3.3.6 では、HBase が他のストレージシステム実装でも実行できるように、HDFS API と先行書き込みログ (WAL) のセマンティクスのサポートも追加されています。詳細については、「Apache Hadoop ドキュメント」のバージョン [3.3.4](#)、[3.3.5](#)、[3.3.6](#) の変更ログを参照してください。

- AWS SDK for Java バージョン 2 のサポート - Amazon EMR 6.15.0 アプリケーションは、アプリケーションが v2 をサポートしている場合は AWS SDK for Java バージョン 1.12.569 または [2.20.160](#) を使用できます。 <https://github.com/aws/aws-sdk-java/tree/1.12.569> AWS SDK for Java 2.x は、バージョン 1.x のコードベースを大幅に書き換えたものです。今回のバージョンは Java 8 以降をベースに構築されており、これまで要望の声が高かった機能が複数追加されています。このバージョンでは、ノンブロッキング I/O がサポートされており、ランタイム時にさまざまな HTTP 実装をプラグインすることができます。SDK for Java v1 から v2 への移行ガイドを含めた詳細については、[AWS SDK for Java バージョン 2 ガイド](#)を参照してください。

## 変更点、機能強化、解決した問題

- 高可用性 EMR クラスターを改善するため、このリリースでは、IPv6 エンドポイントを使用するローカルホスト上の Amazon EMR デモンへの接続が可能になりました。
- このリリースでは、TLS 1.2 が高可用性クラスターのすべてのプライマリノードで ZooKeeper プロビジョニングされた と通信できるようになります。
- このリリースでは、プライマリノードで管理されるトランザクションログファイルの管理 ZooKeeper が改善され、ログファイルが範囲外になり、クラスターオペレーションが中断されるシナリオが最小限に抑えられます。
- このリリースにより、高可用性 EMR クラスターのノード内通信の回復力が高まりました。この改善により、ブートストラップアクションが失敗したり、クラスターの起動に失敗したりする可能性が低くなります。
- Amazon EMR 6.15.0 の Tez では、入カスプリットを Tez グループ分割で非同期的に開くように指定できる設定が導入されています。これにより、1 つの Tez グループ分割で多数の入力分割がある場合に、読み取りクエリのパフォーマンスが向上します。詳細については、「[Tez 非同期スプリットオープン](#)」を参照してください。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>ン)、欧州(フランクフルト)、欧州(チューリッヒ)、欧州(アイルランド)、欧州(ロンドン)、欧州(パリ)、アジアパシフィック(香港)、アジアパシフィック(ムンバイ)、アジアパシフィック(ハイデラバード)、アジアパシフィック(東京)、アジアパシフィック(ソウル)、アジアパシフィック(大阪)、アジアパシフィック(シンガポール)、アジアパシフィック(シドニー)、アジアパシフィック(ジャカルタ)、アジアパシフィック(メルボルン)、アフリカ(ケープタウン)、南米(サンパウロ)、中東(バーレーン)、中東(アラブ首長国連邦)、カナダ(中部)、イスラエル(テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国(北京)、中</p>

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			国 (寧夏)、カナダ西部 (カルガリー)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 13 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

### 6.15.0 のデフォルト Java バージョン

Amazon EMR リリース 6.12.0 以降では、Trino を除き、Amazon Corretto 8 を使用するすべてのアプリケーションをデフォルトでサポートします。Trino の場合、Amazon EMR は Amazon EMR リリース 6.9.0 以降、デフォルトで Amazon Corretto 17 をサポートしています。Amazon EMR は Amazon Corretto 11 および 17 を使用する一部のアプリケーションもサポートしています。これらのアプリケーションのリストを以下の表に示します。クラスターのデフォルト JVM を変更する場合は、クラスター上で実行される各アプリケーションに対して「[特定の Java 仮想マシンを使用するようにアプリケーションを設定](#)」の手順を実行してください。1 つのクラスターに使用できる Java ランタイムバージョンは 1 つだけです。Amazon EMR は、同じクラスター上の異なるランタイムバージョンで異なるノードやアプリケーションを実行することをサポートしていません。

Amazon EMR は Apache Spark、Apache Hadoop、および Apache Hive で Amazon Corretto 11 と 17 の両方をサポートしていますが、これらのバージョンの Corretto を使用すると、一部のワークロードでパフォーマンスが低下する可能性があります。デフォルトを変更する前に、ワークロードをテストすることをお勧めします。

次の表は、Amazon EMR 6.15.0 のアプリケーションのデフォルト Java バージョンを示しています。

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
差分	17、11、8
Flink	11、8
Ganglia	8
HBase	11、8
HCatalog	17、11、8
Hadoop	17、11、8
[Hive]	17、11、8
Hudi	17、11、8
Iceberg	17、11、8
Livy	17、11、8
Oozie	17、11、8
フェニックス	8
PrestoDB	8
Spark	17、11、8
Spark RAPIDS	17、11、8
Sqoop	8
Tez	17、11、8
Trino	17
Zeppelin	8
Pig	8

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
Zookeeper	8

### 6.15.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.4.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>delta-standalone-connectors</code>	0.6.0	Delta コネクタは、Flink、Hive、Presto などのエンジンと Delta Lake を統合するためのさまざまなランタイムを提供します。



コンポーネント	Version	説明
emr-ddb	5.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.8.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.12.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.29.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.8.0	EMR S3Select コネクタ
emr-wal-cli	1.2.0	emrwal のリスト/削除に使用される CLI。
emrfs	2.60.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.17.1-amzn-1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.17.1-amzn-1	Apache Flink の EMR ノード上のリソースの管理 JobManager。

コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.3.6-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.6-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.6-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.6-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.6-amzn-1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.3.6-amzn-1	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	3.3.6-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.3.6-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.6-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.6-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.6-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.17-amzn-3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.17-amzn-3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.17-amzn-3	HBase コマンドラインクライアント。
hbase-rest-server	2.4.17-amzn-3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.17-amzn-3	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hbase-operator-tools	2.4.17-amzn-3	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-8	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-8	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-8	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-8	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-8	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-8	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-8	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.14.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.14.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hudi-trino	0.14.0-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.14.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.4.0-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.5.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。

コンポーネント	Version	説明
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.7.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.3	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.283-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.283-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.283-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	426-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	426-amzn-0	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
trino-client	426-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.4.1-amzn-2	Spark コマンドラインクライアント。
spark-history-server	3.4.1-amzn-2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.4.1-amzn-2	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.4.1-amzn-2	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	23.08.1-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-6	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-6	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

## 6.15.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。



## emr-6.15.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop の KMS java ホームを変更します。	Not available.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.



分類	説明	再設定アクション
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformation	Presto の lakeformation.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.

分類	説明	再設定アクション
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduceHistoryServer.

分類	説明	再設定アクション
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.15.0 変更ログ

### 6.15.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-11-17	ドキュメント公開	Amazon EMR 6.15.0 リリースノートの初版が公開されました
2023-11-17	デプロイ完了	Amazon EMR 6.15.0 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>

日付	イベント	説明
2023-11-13	初回リリース	Amazon EMR 6.15.0 が初回商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.14.0

### 6.14.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
AWS SDK for Java	1.12.543	1.12.513	1.12.490	1.12.446
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-



	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
Delta	2.4.0	2.4.0	2.4.0	2.2.0
Flink	1.17.1	1.17.0	1.17.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.17	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.281	0.281	0.281	0.279
Spark	3.4.1	3.4.1	3.4.0	3.3.2

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	422	414	414	410
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

## 6.14.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.14.0 に関する情報が含まれています。6.13.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.14.0 変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- Amazon EMR 6.14.0 以降で作成したクラスターについて、ap-southeast-3 アジアパシフィック (ジャカルタ) リージョンで [Amazon EMR Managed Scaling](#) を利用できるようになりました。

### 変更点、機能強化、解決した問題

- 6.14.0 リリースでは、Amazon EC2 で実行されている Amazon EMR でのログ管理が最適化されています。その結果、クラスターログのストレージコストがわずかに削減される可能性があります。
- 6.14.0 リリースでは、Amazon EBS ボリュームのサイズに大きなばらつきがあるさまざまなコアインスタンスを考慮して、スケーリングワークフローが改善されています。この改善はコアノードにのみ適用され、タスクノードのスケールダウン操作には影響しません。
- 6.14.0 リリースでは、Amazon EMR が Apache Hadoop YARN ResourceManager and HDFS NameNode などのオープンソースアプリケーションとやり取りする方法が改善されています。こ

の改善により、クラスタースケリングによる運用遅延のリスクが軽減され、オープンソースアプリケーションとの接続の問題が原因で発生する起動障害が軽減されます。

- 6.14.0 リリースでは、クラスター起動時のアプリケーションのインストールが最適化されます。これにより、Amazon EMR アプリケーションの特定の組み合わせのクラスター起動時間が短縮されます。
- 6.14.0 リリースでは、カスタムドメインの VPC で実行されているクラスターがコアノードまたはタスクノードの再起動に遭遇したときに、クラスターのスケールダウン操作が停止することがある問題が修正されています。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィッ

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
-------------------------------------	-------------------------	-------	--------------

ク (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 17 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 9 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

### 6.14.0 のデフォルト Java バージョン

Amazon EMR リリース 6.12.0 以降では、Trino を除き、Amazon Corretto 8 を使用するすべてのアプリケーションをデフォルトでサポートします。Trino の場合、Amazon EMR は Amazon EMR リリース 6.9.0 以降、デフォルトで Amazon Corretto 17 をサポートしています。Amazon EMR は Amazon Corretto 11 および 17 を使用する一部のアプリケーションもサポートしています。これらのアプリケーションのリストを以下の表に示します。クラスターのデフォルト JVM を変更する場合は、クラスター上で実行される各アプリケーションに対して「[特定の Java 仮想マシンを使用するようにアプリケーションを設定](#)」の手順を実行してください。1 つのクラスターに使用できる Java ランタイムバージョンは 1 つだけです。Amazon EMR は、同じクラスター上の異なるランタイムバージョンで異なるノードやアプリケーションを実行することをサポートしていません。

Amazon EMR は Apache Spark、Apache Hadoop、および Apache Hive で Amazon Corretto 11 と 17 の両方をサポートしていますが、これらのバージョンの Corretto を使用すると、一部のワークロードでパフォーマンスが低下する可能性があります。デフォルトを変更する前に、ワークロードをテストすることをお勧めします。

次の表は、Amazon EMR 6.14.0 のアプリケーションのデフォルト Java バージョンを示しています。

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
差分	<b>17</b> 、11、8
Flink	11、 <b>8</b>

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
Ganglia	8
HBase	11、8
HCatalog	17、11、8
Hadoop	17、11、8
[Hive]	17、11、8
Hudi	17、11、8
Iceberg	17、11、8
Livy	17、11、8
Oozie	17、11、8
フェニックス	8
PrestoDB	8
Spark	17、11、8
Spark RAPIDS	17、11、8
Sqoop	8
Tez	17、11、8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

## 6.14.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.4.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>delta-standalone-connectors</code>	0.6.0	Delta コネクタは、Flink、Hive、Presto などのエンジンと Delta Lake を統合するためのさまざまなランタイムを提供します。
<code>emr-ddb</code>	5.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.7.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.11.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.28.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.7.0	EMR S3Select コネクタ
emr-wal-cli	1.1.0	emrwal のリスト/削除に使用される CLI。
emrfs	2.59.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.17.1-amzn-0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.17.1-amzn-0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.3.3-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.3-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.3-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-6	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.3.3-amzn-6	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。



コンポーネント	Version	説明
hadoop-mapred	3.3.3-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-6	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.3-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.17-amzn-2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.17-amzn-2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.17-amzn-2	HBase コマンドラインクライアント。
hbase-rest-server	2.4.17-amzn-2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.17-amzn-2	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.17-amzn-2	Apache HBase クラスター用の修復ツール。

コンポーネント	Version	説明
hcatalog-client	3.1.3-amzn-7	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-7	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-7	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-7	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-7	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-7	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-7	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.13.1-amzn-2	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.13.1-amzn-2	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.13.1-amzn-2	Trino を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hudi-spark	0.13.1-amzn-2	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.3.1-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.5.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
opencv	4.7.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.3	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.281-amzn-2	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.281-amzn-2	いくつかのクエリを実行するサービス。
presto-client	0.281-amzn-2	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	422-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	422-amzn-0	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
trino-client	422-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.4.1-amzn-1	Spark コマンドラインクライアント。
spark-history-server	3.4.1-amzn-1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.4.1-amzn-1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.4.1-amzn-1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	23.06.0-amzn-2	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-5	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-5	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

## 6.14.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-6.14.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.



分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop の KMS java ホームを変更します。	Not available.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformation	Presto の lakeformation.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.

分類	説明	再設定アクション
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduceHistoryServer.

分類	説明	再設定アクション
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.14.0 変更ログ

### 6.14.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-11-02	デプロイ完了	Amazon EMR 6.14.0 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-10-10	ドキュメント公開	Amazon EMR 6.14.0 リリースノートの初版が公開されました

日付	イベント	説明
2023-10-04	初回リリース	Amazon EMR 6.14.0 が初回商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.13.0

### 6.13.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
AWS SDK for Java	1.12.513	1.12.490	1.12.446	1.12.446
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
Delta	2.4.0	2.4.0	2.2.0	2.2.0
Flink	1.17.0	1.17.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.281	0.281	0.279	0.279
Spark	3.4.1	3.4.0	3.3.2	3.3.2

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	414	414	410	410
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

## 6.13.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.13.0 に関する情報が含まれています。6.12.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.13.0 変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

### 変更点、機能強化、解決した問題

- 6.13.0 リリースでは、Amazon EMR ログ管理デーモンが改善され、クラスター終了コマンドが発行されたときに、すべてのログが定期的に Amazon S3 にアップロードされるようになりました。これにより、クラスターの終了が速くなります。
- 6.13.0 リリースでは Amazon EMR のログ管理機能が強化され、すべてのログファイルを Amazon S3 に一貫してタイムリーにアップロードできるようになりました。これは特に、長時間稼働する EMR クラスターにメリットをもたらします。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)</p>

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

### 6.13.0 のデフォルト Java バージョン

Amazon EMR リリース 6.12.0 以降では、Trino を除き、Amazon Corretto 8 を使用するすべてのアプリケーションをデフォルトでサポートします。Trino の場合、Amazon EMR は Amazon EMR リリース 6.9.0 以降、デフォルトで Amazon Corretto 17 をサポートしています。Amazon EMR は Amazon Corretto 11 および 17 を使用する一部のアプリケーションもサポートしています。これらのアプリケーションのリストを以下の表に示します。クラスターのデフォルト JVM を変更する場合は、クラスター上で実行される各アプリケーションに対して「[特定の Java 仮想マシンを使用するようにアプリケーションを設定](#)」の手順を実行してください。1 つのクラスターに使用できる Java ランタイムバージョンは 1 つだけです。Amazon EMR は、同じクラスター上の異なるランタイムバージョンで異なるノードやアプリケーションを実行することをサポートしていません。

Amazon EMR は Apache Spark、Apache Hadoop、および Apache Hive で Amazon Corretto 11 と 17 の両方をサポートしていますが、これらのバージョンの Corretto を使用すると、一部のワークロードでパフォーマンスが低下する可能性があります。デフォルトを変更する前に、ワークロードをテストすることをお勧めします。

次の表は、Amazon EMR 6.13.0 のアプリケーションのデフォルト Java バージョンを示しています。

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
差分	<b>17</b> 、11、8
Flink	11、 <b>8</b>

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
Ganglia	8
HBase	11、8
HCatalog	17、11、8
Hadoop	17、11、8
[Hive]	17、11、8
Hudi	17、11、8
Iceberg	17、11、8
Livy	17、11、8
Oozie	17、11、8
フェニックス	8
PrestoDB	8
Spark	17、11、8
Spark RAPIDS	17、11、8
Sqoop	8
Tez	17、11、8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

### 6.13.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.4.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>delta-standalone-connectors</code>	0.6.0	Delta コネクタは、Flink、Hive、Presto などのエンジンと Delta Lake を統合するためのさまざまなランタイムを提供します。
<code>emr-ddb</code>	5.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.6.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.10.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.27.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.6.0	EMR S3Select コネクタ
emr-wal-cli	1.1.0	emrwal のリスト/削除に使用される CLI。
emrfs	2.58.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.17.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.17.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.3.3-amzn-5	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.3-amzn-5	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.3-amzn-5	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-5	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-5	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.3.3-amzn-5	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-5	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	3.3.3-amzn-5	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-5	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-5	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.3-amzn-5	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.17-amzn-1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.17-amzn-1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.17-amzn-1	HBase コマンドラインクライアント。
hbase-rest-server	2.4.17-amzn-1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.17-amzn-1	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.17-amzn-1	Apache HBase クラスター用の修復ツール。

コンポーネント	Version	説明
hcatalog-client	3.1.3-amzn-6	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-6	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-6	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-6	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-6	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-6	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-6	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.13.1-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.13.1-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.13.1-amzn-1	Trino を Hudi で実行するためのバンドルライブラリ。



コンポーネント	Version	説明
hudi-spark	0.13.1-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.3.0-amzn-1	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.5.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
opencv	4.7.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.3	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.281-amzn-1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.281-amzn-1	いくつかのクエリを実行するサービス。
presto-client	0.281-amzn-1	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	414-amzn-1	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	414-amzn-1	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
trino-client	414-amzn-1	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.4.1-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.4.1-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.4.1-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.4.1-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	23.06.0-amzn-1	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-4	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-4	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

### 6.13.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-6.13.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.



分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop の KMS java ホームを変更します。	Not available.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformation	Presto の lakeformation.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.

分類	説明	再設定アクション
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduceHistoryServer.

分類	説明	再設定アクション
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.13.0 変更ログ

### 6.13.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-09-23	デプロイ完了	Amazon EMR 6.13.0 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-09-12	ドキュメント公開	Amazon EMR 6.13.0 リリースノートの初版が公開されました

日付	イベント	説明
2023-09-01	初回リリース	Amazon EMR 6.13.0 が初回商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.12.0

### 6.12.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
AWS SDK for Java	1.12.490	1.12.446	1.12.446	1.12.397
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
Delta	2.4.0	2.2.0	2.2.0	2.2.0
Flink	1.17.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.15	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.10.0
Iceberg	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.281	0.279	0.279	0.278
Spark	3.4.0	3.3.2	3.3.2	3.3.1

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	414	410	410	403
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

## 6.12.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.12.0 に関する情報が含まれています。6.11.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.12.0 変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- Amazon EMR リリース 6.12.0 以降では、Apache Livy、Apache Hive through HiveServer2 (HS2)、Trino、Presto、Hue との LDAP 統合がサポートされています。6.12.0 以降を使用する EMR クラスターに Apache Spark と Apache Hadoop をインストールし、LDAP を使用するように設定することもできます。詳細については、「[Amazon EMR での認証に Active Directory または LDAP サーバーを使用する](#)」を参照してください。

### 変更点、機能強化、解決した問題

- Amazon EMR リリース 6.12.0 以降では、Java 11 ランタイム環境で Flink を実行できます。詳細については、「[Flink が Java 11 で実行されるよう設定する](#)」を参照してください。
- 6.12.0 リリースでは、Presto または Trino を実行する EMR クラスターのクラスタースケールングワークフローに新しい再試行メカニズムが追加されています。この改善により、サイズ変更操作が

- 1 回失敗したためにクラスターのサイズ変更が無期限に停止するリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。
- 6.12.0 リリースでは、正常に廃止されているコアノードが完全に廃止される前に何らかの理由で異常が発生すると、クラスターのスケールダウン操作が停止することがある問題が修正されています。
  - 6.12.0 リリースでは、クラスターのスケールダウンロジックが改善され、クラスターの HDFS レプリケーション係数設定を下回るコアノードのスケールダウンが試みられることがなくなりました。これはデータの冗長性要件に合致し、スケーリング操作が停止する可能性が低くなります。
  - 6.12.0 リリースでは、インスタンスの状態変化を記録する速度が向上し、Amazon EMR のヘルスマonitoringサービスのパフォーマンスと効率が向上します。この改善により、複数のカスタムクライアントツールやサードパーティーアプリケーションを実行しているクラスターノードのパフォーマンスが低下する可能性が低くなります。
  - 6.12.0 リリースでは、Amazon EMR のクラスター上のログ管理デーモンのパフォーマンスが向上しています。その結果、同時実行性の高いステップを実行する EMR クラスターでは、パフォーマンスが低下する可能性が低くなります。
  - Amazon EMR リリース 6.12.0 では、ログ管理デーモンがアップグレードされ、ローカルインスタンスストレージ上のオープンファイルハンドルでアクティブに使用されているすべてのログと、関連するプロセスが識別されるようになりました。このアップグレードにより、ログが Amazon S3 にアーカイブされた後に Amazon EMR がファイルを適切に削除し、ストレージスペースを再利用できるようになります。
  - 6.12.0 リリースには、ローカルクラスターファイルシステム内の空で未使用のステップディレクトリを削除するログ管理デーモンの機能強化が含まれています。空のディレクトリが多すぎると、Amazon EMR デーモンのパフォーマンスが低下し、ディスクが過剰に使用される可能性があります。
  - 6.12.0 リリースでは、YARN タイムラインサーバーのログのログローテーションが可能になります。これにより、特に長時間稼働するクラスターのディスク過剰使用シナリオが最小限に抑えられます。
  - Amazon EMR 6.10.0 以降では、デフォルトのルートボリュームサイズが 15 GB に増えました。以前のリリースでは、デフォルトのルートボリュームサイズは 10 GB でした。
  - Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国



OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)



## 6.12.0 のデフォルト Java バージョン

Amazon EMR リリース 6.12.0 以降では、Trino を除き、Amazon Corretto 8 を使用するすべてのアプリケーションをデフォルトでサポートします。Trino の場合、Amazon EMR は Amazon EMR リリース 6.9.0 以降、デフォルトで Amazon Corretto 17 をサポートしています。Amazon EMR は Amazon Corretto 11 および 17 を使用する一部のアプリケーションもサポートしています。これらのアプリケーションのリストを以下の表に示します。クラスターのデフォルト JVM を変更する場合は、クラスター上で実行される各アプリケーションに対して「[特定の Java 仮想マシンを使用するようにアプリケーションを設定](#)」の手順を実行してください。1 つのクラスターに使用できる Java ランタイムバージョンは 1 つだけです。Amazon EMR は、同じクラスター上の異なるランタイムバージョンで異なるノードやアプリケーションを実行することをサポートしていません。

Amazon EMR は Apache Spark、Apache Hadoop、および Apache Hive で Amazon Corretto 11 と 17 の両方をサポートしていますが、これらのバージョンの Corretto を使用すると、一部のワークロードでパフォーマンスが低下する可能性があります。デフォルトを変更する前に、ワークロードをテストすることをお勧めします。

次の表は、Amazon EMR 6.12.0 のアプリケーションのデフォルト Java バージョンを示しています。

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
差分	17、11、8
Flink	11、8
Ganglia	8
HBase	11、8
HCatalog	17、11、8
Hadoop	17、11、8
[Hive]	17、11、8
Hudi	17、11、8
Iceberg	17、11、8

アプリケーション	Java/Amazon Corretto バージョン (デフォルトは太字)
Livy	17、11、8
Oozie	17、11、8
フェニックス	8
PrestoDB	8
Spark	17、11、8
Spark RAPIDS	17、11、8
Sqoop	8
Tez	17、11、8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

## 6.12.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオー

プンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.4.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
delta-standalone-connectors	0.6.0	Delta コネクタは、Flink、Hive、Presto などのエンジンと Delta Lake を統合するためのさまざまなランタイムを提供します。
emr-ddb	5.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.5.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.9.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.26.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.5.0	EMR S3Select コネクタ

コンポーネント	Version	説明
emr-wal-cli	1.1.0	emrwal のリスト/削除に使用される CLI。
emrfs	2.57.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.17.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.17.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.3.3-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.3-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	3.3.3-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.3.3-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.3.3-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.3-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	2.4.17-amzn-0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.17-amzn-0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.17-amzn-0	HBase コマンドラインクライアント。
hbase-rest-server	2.4.17-amzn-0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.17-amzn-0	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.17-amzn-0	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-5	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-5	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-5	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-5	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-5	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	3.1.3-amzn-5	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-5	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.13.1-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.13.1-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.13.1-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.13.1-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.3.0-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.7.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.3	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供する軽量サーバー
presto-coordinator	0.281-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。



コンポーネント	Version	説明
presto-worker	0.281-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.281-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	414-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	414-amzn-0	いくつかのクエリを実行するサービス。
trino-client	414-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.4.0-amzn-0	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	3.4.0-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.4.0-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.4.0-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	23.06.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-3	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-3	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

## 6.12.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

### emr-6.12.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode,

分類	説明	再設定アクション
		Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.

分類	説明	再設定アクション
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.



分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop の KMS java ホームを変更します。	Not available.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformation	Presto の lakeformation.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.

分類	説明	再設定アクション
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduceHistoryServer.



分類	説明	再設定アクション
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.12.0 変更ログ

### 6.12.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-07-27	ドキュメントの更新	6.12 の Java オプションを更新し、JVM を更新するための Oozie チュートリアルを追加しました
2023-07-21	デプロイ完了	Amazon EMR 6.12.0 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>

日付	イベント	説明
2023-07-21	ドキュメント公開	Amazon EMR 6.12.0 リリースノートの初版が公開されました
2023-07-12	初回リリース	Amazon EMR 6.12.0 が初回商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.11.1

### 6.11.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
AWS SDK for Java	1.12.446	1.12.446	1.12.397	1.12.397
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.2.0
Flink	1.16.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.5.0	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.279	0.279	0.278	0.278
Spark	3.3.2	3.3.2	3.3.1	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	410	410	403	403
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

### 6.11.1 リリースノート

次のリリースノートには、Amazon EMR リリース 6.11.1 に関する情報が含まれています。6.11.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.11.1 変更ログ](#)」を参照してください。

#### 変更点、機能強化、解決した問題

- ノードが廃止試行と同時に追加または削除されると、ロック競合が原因でデッドロックに陥る可能性があります。その結果、Hadoop Resource Manager (YARN) が応答しなくなり、着信コンテナと現在実行中のコンテナのすべてに影響が及びます。
- このリリースには、高可用性クラスターが再起動後に障害状態から回復できるようにする変更が含まれています。
- このリリースには Hue と HBase のセキュリティ修正が含まれています。
- このリリースでは、Amazon EMR で Spark でワークロードを実行しているクラスターが contains、startsWith、endsWith、および like でメッセージが表示されずに間違った結果を受け取る可能性があるという問題が修正されています。この問題は、Amazon EMR Hive3 メタ

ストアサーバー (HMS) のメタデータを含むパーティション化されたフィールドで式を使用する場合に発生します。

- このリリースでは、ユーザー定義関数 (UDF) がない場合の Glue 側のスロットリングの問題が修正されています。
- このリリースでは、YARN が廃止された場合に、ログプッシャーが S3 にコンテナログをプッシュする前に、ノードログ集約サービスによってコンテナログが削除される問題が修正されています。
- このリリースでは、Hadoop で Node Label が有効になっている場合のス FairShare ケジューラメトリクスの問題が修正されています。
- このリリースでは、spark-defaults.conf の spark.yarn.heterogeneousExecutors.enabled 設定にデフォルトの true 値を設定したときに Spark のパフォーマンスに影響する問題が修正されています。
- このリリースでは、低減タスクがシャッフルデータの読み取りに失敗する問題が修正されています。この問題により、メモリ破損エラーで Hive のクエリが失敗していました。
- このリリースでは、Presto または Trino を実行する EMR クラスターのクラスタースケールリングワークフローに新しい再試行メカニズムが追加されています。この改善により、サイズ変更操作が 1 回失敗したためにクラスターのサイズ変更が無期限に停止するリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。
- このリリースでは、クラスターのスケールダウンロジックが改善され、クラスターの HDFS レプリケーション係数設定を下回るコアノードのスケールダウンが試みられることがなくなりました。これはデータの冗長性要件に合致し、スケールリング操作が停止する可能性が低くなります。
- ログ管理デーモンがアップグレードされ、ローカルインスタンスストレージ上のオープンファイルハンドルでアクティブに使用されているすべてのログと、関連するプロセスが識別されるようになりました。このアップグレードにより、ログが Amazon S3 にアーカイブされた後に Amazon EMR がファイルを適切に削除し、ストレージスペースを再利用できるようになります。
- このリリースには、ローカルクラスターファイルシステム内の空で未使用のステップディレクトリを削除するログ管理デーモンの機能強化が含まれています。空のディレクトリが多すぎると、Amazon EMR デーモンのパフォーマンスが低下し、ディスクが過剰に使用される可能性があります。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)</p>

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 4.14.320 727.0		2023 年 8 月 14 日	連邦)、カナダ (中部)、イスラエル (テルアビブ)  米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

### 6.11.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.2.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>delta-standalone-connectors</code>	0.6.0	Delta コネクタは、Flink、Hive、Presto などのエンジンと Delta Lake を統合するためのさまざまなランタイムを提供します。
<code>emr-ddb</code>	5.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.4.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.8.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.25.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.4.0	EMR S3Select コネクタ
emr-wal-cli	1.1.0	emrwal のリスト/削除に使用される CLI。
emrfs	2.56.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.16.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.16.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.3.3-amzn-3.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.3-amzn-3.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.3-amzn-3.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-3.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-3.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.3.3-amzn-3.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-3.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	3.3.3-amzn-3.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-3.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-3.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.3-amzn-3.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.15-amzn-1.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.15-amzn-1.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.15-amzn-1.1	HBase コマンドラインクライアント。
hbase-rest-server	2.4.15-amzn-1.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.15-amzn-1.1	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.15-amzn-1.1	Apache HBase クラスター用の修復ツール。

コンポーネント	Version	説明
hcatalog-client	3.1.3-amzn-4.1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-4.1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-4.1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-4.1	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-4.1	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-4.1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-4.1	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.13.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.13.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.13.0-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hudi-spark	0.13.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.2.0-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.2	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.279-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.279-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.279-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	410-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	410-amzn-0	いくつかのクエリを実行するサービス。



コンポーネント	Version	説明
trino-client	410-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.3.2-amzn-0.1	Spark コマンドラインクライアント。
spark-history-server	3.3.2-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.3.2-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.3.2-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	23.02.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-2.1	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-2.1	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

### 6.11.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-6.11.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.



分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformat ion	Presto の lakeformation.prop erties ファイルの値を変更し ます。	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の musql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.propert ies ファイルの値を変更しま す。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファ イルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)

分類	説明	再設定アクション
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.



分類	説明	再設定アクション
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.11.1 変更ログ

### 6.11.1 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-08-30	リリースノートの更新	リリースノートにコントロールプレーン関連の修正をいくつか追加しました
2023-08-21	ドキュメント公開	Amazon EMR 6.11.1 リリースノートの初版が公開されました
2023-08-16	デプロイ完了	Amazon EMR 6.11.1 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-08-04	初回リリース	Amazon EMR 6.11.1 が限定商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.11.0

### 6.11.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
AWS SDK for Java	1.12.446	1.12.397	1.12.397	1.12.170
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.1.0
Flink	1.16.0	1.16.0	1.16.0	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.13

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0
Hue	4.11.0	4.10.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.279	0.278	0.278	0.276
Spark	3.3.2	3.3.1	3.3.1	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.10.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
Trino (PrestoSQL)	410	403	403	398
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

## 6.11.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.11.0 に関する情報が含まれています。6.10.0 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR 6.11.0 は、Apache Spark 3.3.2-amzn-0、Apache Spark RAPIDS 23.02.0-amzn-0、CUDA 11.8.0、Apache Hudi 0.13.0-amzn-0、Apache Iceberg 1.2.0-amzn-0、Trino 410-amzn-0、PrestoDB 0.279-amzn-0 をサポートしています。

### 変更点、機能強化、解決した問題

- Amazon EMR 6.11.0 では、DynamoDB コネクタがバージョン 5.0.0 にアップグレードされています。バージョン 5.0.0 では `aws.sdk.java.config` を使用します AWS SDK for Java 2.x。以前のリリースでは AWS SDK for Java 1.x が使用されていました。このアップグレードの結果として、Amazon EMR 6.11 で DynamoDB コネクタを使用する前にコードをテストすることを強くお勧めします。
- Amazon EMR 6.11.0 の DynamoDB コネクタは、DynamoDB サービスを呼び出すとき、`dynamodb.endpoint` プロパティに指定されたリージョンの値を使用します。`dynamodb.endpoint` を使用するとき `dynamodb.region` も設定し、両方のプロパティで同じ AWS リージョンをターゲットにすることをお勧めします。`dynamodb.endpoint` を設定しない場合 `dynamodb.region`、Amazon EMR 6.11.0 の DynamoDB コネクタは無効なリージョン例外を返し、Amazon EC2 インスタンスメタデータサービス (IMDS) からの AWS リージョン 情報の照合を試みます。コネクタが IMDS からリージョンを取得できない場合は、デフォルトで米国東部 (バージニア北部) (`us-east-1`) になります。次のエラーは、`dynamodb.region` プロパティを適切に設定しない場合に発生する可能性のある無効なリージョン例外の例です。2.x error

`software.amazon.awssdk.services.dynamodb.model.DynamoDbException:`

`Credential should be scoped to a valid region.` AWS SDK for Java へのアップグレードの影響を受けるクラスの詳細については、Amazon EMR - DynamoDB コネクタの GitHub リポジトリの「[1.x AWS SDK for Java から 2.x \(#175\) へのアップグレード](#)」コミットを参照してください。

- このリリースでは、列の名前変更操作後に Delta Lake を使用して Amazon S3 に Delta テーブルデータを保存すると列データが NULL になる問題が修正されています。Delta Lake のこの実験的機能の詳細については、「Delta Lake User Guide」の「[Column rename operation](#)」を参照してください。
- 6.11.0 リリースでは、複数のプライマリノードがあるクラスターからプライマリノードの 1 つを複製してエッジノードを作成する際に発生する可能性のある問題が修正されています。複製されたエッジノードは、スケールダウン操作で遅延を引き起こしたり、プライマリノードのメモリ使用率が高くなったりする可能性があります。EMR クラスターと通信するエッジノードを作成する方法の詳細については、の `aws-samples` リポジトリの「[エッジノードの作成](#)」を参照してください GitHub。
- 6.11.0 リリースでは、Amazon EMR が再起動後に Amazon EBS ボリュームをインスタンスに再マウントするために使用する自動化プロセスが改善されました。
- 6.11.0 リリースでは、Amazon EMR が Amazon に発行する Hadoop メトリクスに断続的なギャップが生じる問題が修正されています CloudWatch。
- 6.11.0 リリースでは、クラスターのノードの除外リストを含む YARN 構成ファイルの更新が、ディスクの過剰使用により中断される EMR クラスターの問題が修正されています。更新が不完全だと、今後のクラスターのスケールダウン操作が妨げられます。このリリースでは、クラスターが正常に動作し、スケールアップ操作が期待どおりに機能することが保証されます。
- Amazon EMR 6.10.0 以降では、デフォルトのルートボリュームサイズが 15 GB に増えました。以前のリリースでは、デフォルトのルートボリュームサイズは 10 GB でした。
- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。このアプローチでは、マネージドスケールアップが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。

Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために `yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値が `false` に設定されています。

この修正は YARN-9608 によって発生した問題を解決するものですが、マネージドスケールリングが有効になっているクラスターでのシャッフルデータ損失が原因で Hive ジョブが失敗する可能性があります。このリリースでは、Hive ワークロードにも `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` を設定することで、このリスクを軽減しました。この設定は、Amazon EMR リリース 6.11.0 以降でのみ使用できます。

- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

#### Note

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または [ListReleaseLabels](#) API あるいは [list-release-labels](#) CLI アクションを使用するとき Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
--------------------------------------	-------------------------	-------	--------------

(オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国



OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)

## 6.11.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.2.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>delta-standalone-connectors</code>	0.6.0	Delta コネクタは、Flink、Hive、Presto などのエンジンと Delta Lake を統合するためのさまざまなランタイムを提供します。
<code>emr-ddb</code>	5.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.4.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.8.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.25.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.4.0	EMR S3Select コネクタ
emr-wal-cli	1.1.0	emrwal のリスト/削除に使用される CLI。
emrfs	2.56.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.16.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.16.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。



コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.3.3-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.3-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.3-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-3	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.3.3-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	3.3.3-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.3-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.15-amzn-1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.15-amzn-1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.15-amzn-1	HBase コマンドラインクライアント。
hbase-rest-server	2.4.15-amzn-1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.15-amzn-1	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.15-amzn-1	Apache HBase クラスター用の修復ツール。

コンポーネント	Version	説明
hcatalog-client	3.1.3-amzn-4	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-4	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-4	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-4	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-4	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-4	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-4	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.13.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.13.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.13.0-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hudi-spark	0.13.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.2.0-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.2	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.279-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.279-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.279-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	410-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	410-amzn-0	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
trino-client	410-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.3.2-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.3.2-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.3.2-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.3.2-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	23.02.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-2	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-2	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

### 6.11.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-6.11.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the



分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformat ion	Presto の lakeformation.prop erties ファイルの値を変更し ます。	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の musql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.propert ies ファイルの値を変更しま す。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファ イルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)

分類	説明	再設定アクション
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.

分類	説明	再設定アクション
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.11.0 変更ログ

### 6.11.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-08-21	更新	Hadoop 3.3.3 で発生した問題を修正しました。
2023-07-26	更新	新しい OS リリースラベル 2.0.20230612.0 と 2.0.20230628.0。
2023-06-09	デプロイ完了	Amazon EMR 6.11.0 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-06-09	ドキュメント公開	Amazon EMR 6.11.0 リリースノートの初版が公開されました
2023-06-08	初回リリース	Amazon EMR 6.11.0 が初回商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.10.1

### 6.10.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
AWS SDK for Java	1.12.397	1.12.397	1.12.170	1.12.170
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.1.0	2.1.0
Flink	1.16.0	1.16.0	1.15.2	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.13	2.4.13

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.278	0.278	0.276	0.276
Spark	3.3.1	3.3.1	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.10.0	2.10.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2



	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
Trino (PrestoSQL)	403	403	398	398
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

## 6.10.1 リリースノート

次のリリースノートには、Amazon EMR リリース 6.10.1 に関する情報が含まれています。6.10.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.10.1 変更ログ](#)」を参照してください。

### 変更点、機能強化、解決した問題

- ノードが廃止試行と同時に追加または削除されると、ロック競合が原因でデッドロックに陥る可能性があります。その結果、Hadoop Resource Manager (YARN) が応答しなくなり、着信コンテナと現在実行中のコンテナのすべてに影響が及びます。
- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。このアプローチでは、マネージドスケールリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。

Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために `yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値が `false` に設定されています。

この修正は YARN-9608 によって発生した問題を解決するものですが、マネージドスケールリングが有効になっているクラスターでのシャッフルデータ損失が原因で Hive ジョブが失敗する可能性があります。このリリースでは、Hive ワークロードにも `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` を設定することで、このリスクを軽減しました。この設定は、Amazon EMR リリース 6.11.0 以降でのみ使用できます。

- インスタンスグループが設定されているクラスター内のプライマリノードのフェイルオーバー後、メトリクスコレクターはコントロールプレーンにメトリクスを送信しません。

- このリリースには、高可用性クラスターが再起動後に障害状態から回復できるようにする変更が含まれています。
- このリリースには Hue と HBase のセキュリティ修正が含まれています。
- このリリースでは、Amazon EMR で Spark でワークロードを実行しているクラスターが `contains`、`startsWith`、`endsWith`、および `like` でメッセージが表示されずに間違った結果を受け取る可能性があるという問題が修正されています。この問題は、Amazon EMR Hive3 メタストアサーバー (HMS) のメタデータを含むパーティション化されたフィールドで式を使用する場合に発生します。
- このリリースでは、ユーザー定義関数 (UDF) がない場合の Glue 側のスロットリングの問題が修正されています。
- このリリースでは、YARN が廃止された場合に、ログプッシャーが S3 にコンテナログをプッシュする前に、ノードログ集約サービスによってコンテナログが削除される問題が修正されています。
- このリリースでは、Hadoop で Node Label が有効になっている場合のス FairShare ケジューラメトリクスの問題が修正されています。
- このリリースでは、`spark-defaults.conf` の `spark.yarn.heterogeneousExecutors.enabled` 設定にデフォルトの `true` 値を設定したときに Spark のパフォーマンスに影響する問題が修正されています。
- このリリースでは、低減タスクがシャッフルデータの読み取りに失敗する問題が修正されています。この問題により、メモリ破損エラーで Hive のクエリが失敗していました。
- このリリースでは、Presto または Trino を実行する EMR クラスターのクラスタースケールアップワークフローに新しい再試行メカニズムが追加されています。この改善により、サイズ変更操作が 1 回失敗したためにクラスターのサイズ変更が無期限に停止するリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。
- このリリースでは、クラスターのスケールダウンロジックが改善され、クラスターの HDFS レプリケーション係数設定を下回るコアノードのスケールダウンが試みられることがなくなりました。これはデータの冗長性要件に合致し、スケールアップ操作が停止する可能性が低くなります。
- ログ管理デーモンがアップグレードされ、ローカルインスタンスストレージ上のオープンファイルハンドルでアクティブに使用されているすべてのログと、関連するプロセスが識別されるようになりました。このアップグレードにより、ログが Amazon S3 にアーカイブされた後に Amazon EMR がファイルを適切に削除し、ストレージスペースを再利用できるようになります。
- このリリースには、ローカルクラスターファイルシステム内の空で未使用のステップディレクトリを削除するログ管理デーモンの機能強化が含まれています。空のディレクトリが多すぎると、Amazon EMR デーモンのパフォーマンスが低下し、ディスクが過剰に使用される可能性があります。

- このリリースでは、複数のプライマリノードがあるクラスターからプライマリノードの1つを複製してエッジノードを作成する際に発生する可能性のある問題が修正されています。複製されたエッジノードは、スケールダウン操作で遅延を引き起こしたり、プライマリノードのメモリ使用率が高くなったりする可能性があります。EMR クラスターと通信するエッジノードを作成する方法の詳細については、のaws-samplesリポジトリの「[エッジノードの作成](#)」を参照してくださいGitHub。
- このリリースでは、Amazon EMR が再起動後に Amazon EBS ボリュームをインスタンスに再マウントするために使用する自動化プロセスが改善されました。
- このリリースでは、Amazon EMR が Amazon に発行する Hadoop メトリクスに断続的なギャップが生じる問題が修正されています CloudWatch。
- このリリースでは、クラスターのノードの除外リストを含む YARN 構成ファイルの更新が、ディスクの過剰使用により中断される EMR クラスターの問題が修正されています。更新が不完全だと、今後のクラスターのスケールダウン操作が妨げられます。このリリースでは、クラスターが正常に動作し、スケーリング操作が期待どおりに機能することが保証されます。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッ

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>ヒ)、欧州(アイルランド)、欧州(ロンドン)、欧州(パリ)、アジアパシフィック(香港)、アジアパシフィック(ムンバイ)、アジアパシフィック(ハイデラバード)、アジアパシフィック(東京)、アジアパシフィック(ソウル)、アジアパシフィック(大阪)、アジアパシフィック(シンガポール)、アジアパシフィック(シドニー)、アジアパシフィック(ジャカルタ)、アジアパシフィック(メルボルン)、アフリカ(ケープタウン)、南米(サンパウロ)、中東(バーレーン)、中東(アラブ首長国連邦)、カナダ(中部)、イスラエル(テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)、カナダ西部(カルガリー)</p>

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 4.14.320 727.0		2023 年 8 月 14 日	連邦)、カナダ (中部)、イスラエル (テルアビブ)  米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

## 6.10.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.2.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.7.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-notebook-env</code>	1.7.0	Jupyter エンタープライズゲートウェイを含む <code>emr</code> ノートブック用 Conda env

コンポーネント	Version	説明
emr-s3-dist-cp	2.24.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.3.0	EMR S3Select コネクタ
emr-wal-cli	1.0.0	emrwal のリスト/削除に使用される CLI。
emrfs	2.55.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.16.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.16.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。



コンポーネント	Version	説明
hadoop-client	3.3.3-amzn-2.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.3-amzn-2.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.3-amzn-2.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-2.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-2.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httfs-server	3.3.3-amzn-2.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-2.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.3.3-amzn-2.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-2.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-2.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	3.3.3-amzn-2.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.15-amzn-0.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.15-amzn-0.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.15-amzn-0.1	HBase コマンドラインクライアント。
hbase-rest-server	2.4.15-amzn-0.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.15-amzn-0.1	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.15-amzn-0.1	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-3.1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-3.1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-3.1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。

コンポーネント	Version	説明
hive-client	3.1.3-amzn-3.1	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-3.1	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-3.1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-3.1	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.12.2-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.12.2-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.12.2-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.12.2-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.1.0-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。

コンポーネント	Version	説明
jupyterhub	1.5.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	6.0.0-SNAPSHOT	Spark-3 用 Apache Phoenix-Connectors

コンポーネント	Version	説明
phoenix-query-server	6.0.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.278.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.278.1-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.278.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	403-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	403-amzn-0	いくつかのクエリを実行するサービス。
trino-client	403-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト

コンポーネント	Version	説明
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.3.1-amzn-0.1	Spark コマンドラインクライアント。
spark-history-server	3.3.1-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.3.1-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.3.1-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	22.12.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-1.1	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-1.1	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

### 6.10.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-6.10.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.



分類	説明	再設定アクション
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

分類	説明	再設定アクション
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

分類	説明	再設定アクション
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.

分類	説明	再設定アクション
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.

分類	説明	再設定アクション
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.

分類	説明	再設定アクション
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformat ion	Presto の lakeformation.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)



分類	説明	再設定アクション
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)

分類	説明	再設定アクション
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.10.1 変更ログ

### 6.10.1 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-08-30	リリースノートの更新	リリースノートにコントロールプレーン関連の修正をいくつか追加しました
2023-08-21	ドキュメント公開	Amazon EMR 6.10.1 リリースノートの初版が公開されました
2023-08-16	デプロイ完了	Amazon EMR 6.10.1 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-08-04	初回リリース	Amazon EMR 6.10.1 が限定商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.10.0

### 6.10.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
AWS SDK for Java	1.12.397	1.12.170	1.12.170	1.12.170
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.1.0	2.1.0	-
Flink	1.16.0	1.15.2	1.15.2	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.13	2.4.13	2.4.12
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.1.0
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.278	0.276	0.276	0.273
Spark	3.3.1	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.10.0	2.10.0	2.9.1
Tez	0.10.2	0.10.2	0.10.2	0.9.2
Trino (PrestoSQL)	403	398	398	388
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

## 6.10.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.10.0 に関する情報が含まれています。6.9.0 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。



## 新機能

- Amazon EMR 6.10.0 は、Apache Spark 3.3.1、Apache Spark RAPIDS 22.12.0、CUDA 11.8.0、Apache Hudi 0.12.2-amzn-0、Apache Iceberg 1.1.0-amzn-0、Trino 403、PrestoDB 0.278.1 をサポートしています。
- Amazon EMR 6.10.0 には、Hudi テーブル内のデータへの読み取りアクセスを提供するネイティブ Trino-Hudi コネクタが含まれています。コネクタは `trino-cli --catalog hudi` で有効化でき、`trino-connector-hudi` との要件に合わせてコネクタを設定できます。Amazon EMR とのネイティブ統合により、Hudi テーブルのクエリに `trino-connector-hive` を使用する必要がなくなりました。新しいコネクタでサポートされる設定のリストについては、Trino のドキュメントの「[Hudi connector](#)」ページを参照してください。
- Amazon EMR リリース 6.10.0 以降では、Apache Zeppelin と Apache Flink の統合がサポートされています。詳細については、「[Amazon EMR の Zeppelin から Flink ジョブを操作する](#)」を参照してください。

## 既知の問題

- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。このアプローチでは、マネージドスケーリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。

Amazon EMR 6.10.0 でこの問題を回避するには、`yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値を `false` に設定します。Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために設定がデフォルトで `false` にセットされています。

## 変更点、機能強化、解決した問題

- Amazon EMR 6.10.0 では、[Apache Spark 用の Amazon Redshift の統合](#)のために `minimal-json.jar` の依存関係が削除され、必要な Spark-Redshift 関連の jar が Spark のエグゼキュータークラスパス `spark-redshift.jar`、`spark-avro.jar`、および `RedshiftJDBC.jar` に自動的に追加されます。

- 6.10.0 リリースでは、クラスター上のログ管理デーモンが改善され、EMR クラスター内の追加のログフォルダを監視できるようになりました。この改善により、ディスクの過剰使用シナリオが最小限に抑えられます。
- 6.10.0 リリースでは、クラスター上のログ管理デーモンが停止すると、自動的に再起動されます。この改善により、ディスクの過剰使用が原因でノードが異常に見えるリスクが軽減されます。
- Amazon EMR 6.10.0 は EMRFS ユーザーマッピング用のリージョンのエンドポイントをサポートしています。
- Amazon EMR 6.10.0 以降では、デフォルトのルートボリュームサイズが 15 GB に増えました。以前のリリースでは、デフォルトのルートボリュームサイズは 10 GB でした。
- 6.10.0 リリースでは、残りのすべての Spark エグゼキューターが YARN リソースマネージャーを備えた廃止予定のホスト上にあるとき Spark ジョブが停止する問題が修正されています。
- Amazon EMR 6.6.0 から 6.9.x では、動的パーティションと ORDER BY 句または SORT BY 句を使用した INSERT クエリには常に 2 つのリデューサーがあります。この問題は、OSS が変更された [HIVE-20703](#) が原因です。これにより、動的ソートパーティションの最適化がコストベースの決定下に置かれます。ワークロードで動的パーティションのソートが不要な場合は、hive.optimize.sort.dynamic.partition.threshold プロパティを -1 に設定して新機能を無効にし、リデューサーの数を正しく計算することをお勧めします。この問題は、[HIVE-22269](#) の一部として OSS Hive で修正され、Amazon EMR 6.10.0 で修正されています。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

#### Note

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または [ListReleaseLabels](#) API あるいは [list-release-labels](#) CLI アクションを使用するとき Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (UAE)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (UAE)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 207.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (UAE)

## 6.10.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.2.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.7.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-notebook-env</code>	1.7.0	Jupyter エンタープライズゲートウェイを含む <code>emr</code> ノートブック用 Conda env

コンポーネント	Version	説明
emr-s3-dist-cp	2.24.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.3.0	EMR S3Select コネクタ
emr-wal-cli	1.0.0	emrwal のリスト/削除に使用される CLI。
emrfs	2.55.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.16.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.16.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。



コンポーネント	Version	説明
hadoop-client	3.3.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.3.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-2	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.3.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.3.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	3.3.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.15-amzn-0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.15-amzn-0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.15-amzn-0	HBase コマンドラインクライアント。
hbase-rest-server	2.4.15-amzn-0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.15-amzn-0	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.15-amzn-0	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-3	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-3	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-3	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。

コンポーネント	Version	説明
hive-client	3.1.3-amzn-3	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-3	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-3	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-3	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.12.2-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.12.2-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.12.2-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.12.2-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	1.1.0-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。

コンポーネント	Version	説明
jupyterhub	1.5.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.8.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	6.0.0-SNAPSHOT	Spark-3 用 Apache Phoenix-Connectors

コンポーネント	Version	説明
phoenix-query-server	6.0.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.278.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.278.1-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.278.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	403-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	403-amzn-0	いくつかのクエリを実行するサービス。
trino-client	403-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト

コンポーネント	Version	説明
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.3.1-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.3.1-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.3.1-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.3.1-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	22.12.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-1	tez YARN アプリケーションおよびライブラリ。
tez-on-worker	0.10.2-amzn-1	ワーカーノード用 tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

### 6.10.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-6.10.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.



分類	説明	再設定アクション
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

分類	説明	再設定アクション
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

分類	説明	再設定アクション
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.

分類	説明	再設定アクション
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.

分類	説明	再設定アクション
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.

分類	説明	再設定アクション
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformat ion	Presto の lakeformation.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)



分類	説明	再設定アクション
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)

分類	説明	再設定アクション
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino の hudi.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS の kms-logback.xml ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-orahome-site	Sqoop OraOop の orahome-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeper の zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper の log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.10.0 変更ログ

### 6.10.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-08-21	更新	Hadoop 3.3.3 で発生した既知の問題が追加されました。
2023-07-26	更新	新しい OS リリースラベル 2.0.20230612.0 と 2.0.20230628.0。
2023-03-02	デプロイ完了	Amazon EMR 6.10 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-03-02	ドキュメント公開	Amazon EMR 6.10 リリースノートの初版が公開されました
2023-02-27	初回リリース	Amazon EMR 6.10 が限定商用リージョンにデプロイされました

## Amazon EMR リリース 6.9.1

### 6.9.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	2.1.0	-	-
Flink	1.15.2	1.15.2	1.15.1	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.13	2.4.12	2.4.12
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.276	0.276	0.273	0.273
Spark	3.3.0	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.10.0	2.9.1	2.9.1
Tez	0.10.2	0.10.2	0.9.2	0.9.2
Trino (PrestoSQL)	398	398	388	388
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

## 6.9.1 リリースノート

次のリリースノートには、Amazon EMR リリース 6.9.1 に関する情報が含まれています。6.9.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.9.1 変更ログ](#)」を参照してください。



## 変更点、機能強化、解決した問題

- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。このアプローチでは、マネージドスケールリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。

Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために `yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値が `false` に設定されています。

この修正は YARN-9608 によって発生した問題を解決するものですが、マネージドスケールリングが有効になっているクラスターでのシャッフルデータ損失が原因で Hive ジョブが失敗する可能性があります。このリリースでは、Hive ワークロードにも `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` を設定することで、このリスクを軽減しました。この設定は、Amazon EMR リリース 6.11.0 以降でのみ使用できます。

- インスタンスグループが設定されているクラスター内のプライマリノードのフェイルオーバー後、メトリクスコレクターはコントロールプレーンにメトリクスを送信しません。
- このリリースでは、メトリクスコレクターのエンドポイントへの HTTP リクエストが失敗しても再試行する必要がなくなりました。
- このリリースには、高可用性クラスターが再起動後に障害状態から回復できるようにする変更が含まれています。
- このリリースでは、ユーザーが作成した大きな UID がオーバーフロー例外を引き起こしていた問題が修正されています。
- このリリースでは、Amazon EMR 再構成プロセスのタイムアウトの問題が修正されています。
- このリリースにはセキュリティ修正が含まれています。
- このリリースでは、Amazon EMR で Spark でワークロードを実行しているクラスターが `contains`、`startsWith`、`endsWith`、および `like` でメッセージが表示されずに間違った結果を受け取る可能性があるという問題が修正されています。この問題は、Amazon EMR Hive3 メタストアサーバー (HMS) のメタデータを含むパーティション化されたフィールドで式を使用する場合に発生します。
- Amazon EMR 6.6.0 から 6.9.x では、動的パーティションと ORDER BY 句または SORT BY 句を使用した INSERT クエリには常に 2 つのリデューサーがあります。この問題は、OSS が変更された [HIVE-20703](#) が原因です。これにより、動的ソートパーティションの最適化がコ

ストベースの決定下に置かれます。ワークロードで動的パーティションのソートが不要な場合は、`hive.optimize.sort.dynamic.partition.threshold` プロパティを `-1` に設定して新機能を無効にし、リデューサーの数を正しく計算することをお勧めします。この問題は、[HIVE-22269](#) の一部として OSS Hive で修正され、Amazon EMR 6.10.0 で修正されています。

- HDFS をスクラッチディレクトリとして使用し、小さいファイルのマージを有効にしている、テーブルに静的なパーティションパスが含まれている場合、Hive でデータが失われる可能性があります。
- このリリースでは、ETL ジョブの終了時に小さいファイルのマージ (デフォルトでは無効) が有効になっている場合の Hive のパフォーマンスの問題が修正されています。
- このリリースでは、ユーザー定義関数 (UDF) がない場合の Glue 側のスロットリングの問題が修正されています。
- このリリースでは、YARN が廃止された場合に、ログプッシャーが S3 にコンテナログをプッシュする前に、ノードログ集約サービスによってコンテナログが削除される問題が修正されています。
- このリリースでは、HBase の永続的なストアファイル追跡による圧縮/アーカイブファイルの処理が修正されています。
- このリリースでは、`spark-defaults.conf` の `spark.yarn.heterogeneousExecutors.enabled` 設定にデフォルトの `true` 値を設定したときに Spark のパフォーマンスに影響する問題が修正されています。
- このリリースでは、低減タスクがシャッフルデータの読み取りに失敗する問題が修正されています。この問題により、メモリ破損エラーで Hive のクエリが失敗していました。
- このリリースでは、ノード交換中に HDFS NameNode (NN) サービスがセーフモードで停止した場合にノードプロビジョナーが失敗する問題が修正されています。
- このリリースでは、Presto または Trino を実行する EMR クラスターのクラスタースケリングワークフローに新しい再試行メカニズムが追加されています。この改善により、サイズ変更操作が 1 回失敗したためにクラスターのサイズ変更が無期限に停止するリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。
- このリリースでは、クラスターのスケールダウンロジックが改善され、クラスターの HDFS レプリケーション係数設定を下回るコアノードのスケールダウンが試みられることがなくなりました。これはデータの冗長性要件に合致し、スケリング操作が停止する可能性が低くなります。
- ログ管理デーモンがアップグレードされ、ローカルインスタンスストレージ上のオープンファイルハンドルでアクティブに使用されているすべてのログと、関連するプロセスが識別されるようになりました。このアップグレードにより、ログが Amazon S3 にアーカイブされた後に Amazon EMR がファイルを適切に削除し、ストレージスペースを再利用できるようになります。

- このリリースには、ローカルクラスターファイルシステム内の空で未使用のステップディレクトリを削除するログ管理デーモンの機能強化が含まれています。空のディレクトリが多すぎると、Amazon EMR デーモンのパフォーマンスが低下し、ディスクが過剰に使用される可能性があります。
- このリリースでは、複数のプライマリノードがあるクラスターからプライマリノードの1つを複製してエッジノードを作成する際に発生する可能性のある問題が修正されています。複製されたエッジノードは、スケールダウン操作で遅延を引き起こしたり、プライマリノードのメモリ使用率が高くなったりする可能性があります。EMR クラスターと通信するエッジノードを作成する方法の詳細については、のaws-samplesリポジトリの「[エッジノードの作成](#)」を参照してください GitHub。
- このリリースでは、Amazon EMR が再起動後に Amazon EBS ボリュームをインスタンスに再マウントするために使用する自動化プロセスが改善されました。
- このリリースでは、Amazon EMR が Amazon に発行する Hadoop メトリクスに断続的なギャップが生じる問題が修正されています CloudWatch。
- このリリースでは、クラスターのノードの除外リストを含む YARN 構成ファイルの更新が、ディスクの過剰使用により中断される EMR クラスターの問題が修正されています。更新が不完全だと、今後のクラスターのスケールダウン操作が妨げられます。このリリースでは、クラスターが正常に動作し、スケールアップ操作が期待どおりに機能することが保証されます。
- このリリースでは、クラスター上のログ管理デーモンが改善され、EMR クラスター内の追加のログフォルダを監視できるようになりました。この改善により、ディスクの過剰使用シナリオが最小限に抑えられます。
- このリリースでは、クラスター上のログ管理デーモンが停止すると、自動的に再起動されます。この改善により、ディスクの過剰使用が原因でノードが異常に見えるリスクが軽減されます。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			GovCloud (米国西部)、AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)、カナダ西部(カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)、カナダ西部(カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)、カナダ西部(カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)、カナダ西部(カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

## 6.9.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.1.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.6.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-notebook-env</code>	1.7.0	Jupyter エンタープライズゲートウェイを含む <code>emr</code> ノートブック用 Conda env

コンポーネント	Version	説明
emr-s3-dist-cp	2.23.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.2.0	EMR S3Select コネクタ
emrfs	2.54.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.15.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.15.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.3.3-amzn-1.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。

コンポーネント	Version	説明
hadoop-hdfs-datanode	3.3.3-amzn-1.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.3-amzn-1.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-1.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-1.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-https-server	3.3.3-amzn-1.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-1.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.3.3-amzn-1.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-1.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-1.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.3-amzn-1.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	2.4.13-amzn-0.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.13-amzn-0.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.13-amzn-0.1	HBase コマンドラインクライアント。
hbase-rest-server	2.4.13-amzn-0.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.13-amzn-0.1	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.13-amzn-0.1	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-2.1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-2.1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-2.1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-2.1	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-2.1	Hive-hbase クライアント。



コンポーネント	Version	説明
hive-metastore-server	3.1.3-amzn-2.1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-2.1	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.12.1-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.12.1-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.12.1-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.12.1-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	0.14.1-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.7.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	6.0.0-SNAPSHOT	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	6.0.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供する軽量サーバー
presto-coordinator	0.276-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.276-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.276-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	398-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	398-amzn-0	いくつかのクエリを実行するサービス。
trino-client	398-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.3.0-amzn-1.1	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	3.3.0-amzn-1.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.3.0-amzn-1.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.3.0-amzn-1.1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	22.08.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.10.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-0.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

### 6.9.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-6.9.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ

分類	説明	再設定アクション
		er, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.



分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformat ion	Presto の lakeformation.prop erties ファイルの値を変更し ます。	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の musql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.propert ies ファイルの値を変更しま す。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファ イルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)

分類	説明	再設定アクション
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.

分類	説明	再設定アクション
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.

分類	説明	再設定アクション
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.9.1 変更ログ

### 6.9.1 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-08-30	リリースノートの更新	リリースノートにコントロールプレーン関連の修正をいくつか追加しました
2023-08-21	ドキュメント公開	Amazon EMR 6.9.1 リリースノートの初版が公開されました
2023-08-16	デプロイ完了	Amazon EMR 6.9.1 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-08-04	初回リリース	Amazon EMR 6.9.1 が限定商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.9.0

### 6.9.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	-	-	-
Flink	1.15.2	1.15.1	1.15.1	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.12	2.4.12	2.4.4
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
Hudi	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.276	0.273	0.273	0.272
Spark	3.3.0	3.3.0	3.3.0	3.2.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.9.1	2.9.1	2.4.1
Tez	0.10.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	398	388	388	378
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.7

## 6.9.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.9.0 に関する情報が含まれています。Amazon EMR リリース 6.8.0 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR リリース 6.9.0 は、Apache Spark RAPIDS 22.08.0、Apache Hudi 0.12.1、Apache Iceberg 0.14.1、Trino 398、Tez 0.10.2 をサポートしています。
- Amazon EMR リリース 6.9.0 には、新しいオープンソースアプリケーションである [Delta Lake](#) 2.1.0 が含まれています。
- Apache Spark 用の Amazon Redshift の統合は、Amazon EMR リリース 6.9.0 以降に含まれています。以前はオープンソースツールであったこのネイティブインテグレーションは Spark コネクタと呼ばれるもので、これを使用して Apache Spark アプリケーションを構築することで、Amazon Redshift と Amazon Redshift Serverless 内のデータを読み書きできます。詳細については、「[Amazon EMR での Apache Spark 用の Amazon Redshift インテグレーションの使用](#)」を参照してください。
- Amazon EMR リリース 6.9.0 では、クラスターのスケールダウン中に Amazon S3 にログをアーカイブするためのサポートが追加されています。以前は、Amazon S3 にログファイルをアーカイブできるのはクラスター終了時のみでした。この新機能により、クラスターで生成されたログファイルは、ノードが終了した後も Amazon S3 に残ります。詳細については、「[クラスターのログ記録とデバッグを設定する](#)」を参照してください。
- 長時間実行されるクエリをサポートするため、Trino には耐障害性実行メカニズムが組み込まれました。耐障害性実行では、失敗したクエリやそのコンポーネントタスクを再試行することで、クエリの失敗を軽減します。詳細については、「[Trino での耐障害性実行](#)」を参照してください。
- Amazon EMR で Apache Flink を使用して、Apache Hive テーブル、または Iceberg、Kinesis、Kafka などの任意の Flink テーブルソースのメタデータの BATCH 処理と STREAM 処理を統合することができます。、または Amazon EMR API を使用して AWS Management Console AWS CLI、AWS Glue Data Catalog を Flink のメタストアとして指定できます。詳細については、「[Amazon EMR で Flink を設定する](#)」を参照してください。
- Amazon SageMaker Studio を使用して EC2 クラスター上の Amazon EMR で Apache Spark、Apache Hive、Presto クエリに AWS Identity and Access Management (IAM) ランタイムロールと AWS Lake Formation ベースのアクセスコントロールを指定できるようになりました。詳細については、「[Configure runtime roles for Amazon EMR steps](#)」を参照してください。

## 既知の問題

- Amazon EMR リリース 6.9.0 では、Apache Ranger が有効になっているクラスターでは Trino は動作しません。Ranger で Trino を使用する必要がある場合は、[AWS Support](#) にお問い合わせください。
- Amazon Redshift integration for Apache Spark を使用している場合に、time、timetz、timestamp、timestampz のいずれかにマイクロ秒の精度を Parquet 形式で設定していると、コネクタがその時間値を最も近いミリ秒値に四捨五入します。回避策として、テキストアンロード形式 unload\_s3\_format パラメータを使用してください。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、s3://bucket/table/p=a は s3://bucket/table/p=a b のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の

spark.sql.sources.fastS3PartitionDiscovery.enabled 設定を false にセットします。

- Amazon SageMaker Studio から Amazon EMR クラスターへの接続は、403 Forbidden レスポンスコードで断続的に失敗することがあります。このエラーは、クラスターでの IAM ロールの設定に 60 秒以上かかる場合に発生します。回避策として、Amazon EMR パッチをインストールして再試行を有効にし、タイムアウトを最低 300 秒に増やすことができます。以下の手順を使用して、クラスターの起動時にブートストラップアクションを適用します。

1. 次の Amazon S3 URI から、ブートストラップスクリプトと RPM ファイルをダウンロードします。

```
s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
```

```
s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
```

2. 前のステップのファイルを、所有する Amazon S3 バケットにアップロードします。バケットは、クラスターを起動する AWS リージョン 予定のと同じにある必要があります。
3. EMR クラスターを起動するときに、次のブートストラップアクションを含めません。 *bootstrap\_URI* と *RPM\_URI* を Amazon S3 の対応する URI に置き換えます。

```
--bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
```

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、Log4j2 プロパティのファイル名パターン設定が正しくないため、SecretAgent および RecordServer サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- Apache Flink にはネイティブ S3 FileSystem および Hadoop FileSystem コネクタが用意されており、アプリケーションは を作成し、データを Amazon S3 に FileSink 書き込むことができます。これは、次の 2 つの例外のいずれかで FileSink 失敗します。

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError: org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/
```



```
apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/
RetryPolicy;

                                at
  org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-
yarn-common-3.3.3-amzn-0.jar:?]
```

回避策として、Flink の上記の問題を修正する Amazon EMR パッチをインストールできます。クラスターを起動するときにはブートストラップアクションを適用するには、以下の手順を実行します。

1. flink-rpm を Amazon S3 バケットにダウンロードします。RPM パスは `s3://DOC-EXAMPLE-BUCKET/rpms/flink/` です。
2. 次の URI を使用して Amazon S3 からブートストラップスクリプトと RPM ファイルをダウンロードします。を、クラスターを起動する AWS リージョン 予定の `regionName` に置き換えます。

```
s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-
rpms.sh
```

3. Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。Amazon EMR 6.8.0 および 6.9.0 の場合、この方法では、マネージドスケーリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性があります。

[Amazon EMR 6.10.0](#) では、`yarn-site.xml` の

`yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値を `false` に設定することでこの問題を回避できます。Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために設定がデフォルトで `false` にセットされています。

## 変更、拡張、解決した問題

- Amazon EMR リリース 6.9.0 以降では、Amazon EMR によってインストールされ、Log4j ライブラリを使用するすべてのコンポーネントは Log4j バージョン 2.17.1 以降を使用します。
- Spark on Amazon EMR バージョン 6.6.0、6.7.0、6.8.0 で DynamoDB コネクタを使用すると、テーブルから何を読み込んでも空の結果が返されます。この状況は、入力分割が空でないデータを

参照している場合でも変わりません。Amazon EMR リリース 6.9.0 では、この問題が修正されています。

- Amazon EMR 6.9.0 では、Spark SQL を使用してデータを読み取る際の Apache Hudi による Lake Formation ベースのアクセスコントロールに対する限定的なサポートが追加されています。Spark SQL を使用する SELECT クエリがサポートされます。サポートは、列レベルのアクセス制御に限定されます。詳細については、「[Hudi と Lake Formation](#)」を参照してください。
- Amazon EMR 6.9.0 を使用して、[ノードラベル](#)を有効にした Hadoop クラスターを作成すると、[YARN メトリクス API](#) はデフォルトのパーティションではなく、すべてのパーティションの集計情報を返します。詳細については、「[YARN-11414](#)」を参照してください。
- Amazon EMR リリース 6.9.0 では、Trino を、Java 17 を使用するバージョン 398 にアップデートしました。Amazon EMR 6.8.0 で以前サポートされていた Trino のバージョンは Java 11 で動作する Trino 388 でした。この変更の詳細については、Trino のブログの「[Trino updates to Java 17](#)」を参照してください。
- このリリースでは、Apache BigTop と EC2 クラスターの Amazon EMR の起動シーケンス間のタイミングシーケンスの不一致の問題が修正されています。このタイミングシーケンスの不一致は、システムが 2 つ以上の操作を適切なシーケンスで実行するのではなく、同時に実行しようとした場合に発生します。その結果、特定のクラスター構成でインスタンスの起動がタイムアウトし、クラスターの起動時間が遅くなりました。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

#### Note

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または [ListReleaseLabels](#) API あるいは [list-release-labels](#) CLI アクションを使用するとき Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202307.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 1 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



## 6.9.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>delta</code>	2.1.0	Delta Lake は、膨大な分析データセット用のオープンテーブル形式です。
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.6.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-notebook-env</code>	1.7.0	Jupyter エンタープライズゲートウェイを含む <code>emr</code> ノートブック用 Conda env

コンポーネント	Version	説明
emr-s3-dist-cp	2.23.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.2.0	EMR S3Select コネクタ
emrfs	2.54.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.15.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.15.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.3.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。

コンポーネント	Version	説明
hadoop-hdfs-datanode	3.3.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.3.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.3.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.3.3-amzn-1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-https-server	3.3.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.3.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.3.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.3.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.3.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.3.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	2.4.13-amzn-0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.13-amzn-0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.13-amzn-0	HBase コマンドラインクライアント。
hbase-rest-server	2.4.13-amzn-0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.13-amzn-0	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.13-amzn-0	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-2	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-2	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	3.1.3-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.12.1-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.12.1-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.12.1-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.12.1-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	0.14.1-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.7.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	6.0.0-SNAPSHOT	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	6.0.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供する軽量サーバー
presto-coordinator	0.276-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.276-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.276-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	398-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	398-amzn-0	いくつかのクエリを実行するサービス。
trino-client	398-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.3.0-amzn-1	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	3.3.0-amzn-1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.3.0-amzn-1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.3.0-amzn-1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	22.08.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.10.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.10.2-amzn-0	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。



コンポーネント	Version	説明
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

## 6.9.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

### emr-6.9.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ

分類	説明	再設定アクション
		er, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
delta-defaults	Delta の delta-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j2	Livy の log4j2.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-lakeformation	Presto の lakeformation.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-delta	Trino の delta.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino の exchange-manager.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)

分類	説明	再設定アクション
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.

分類	説明	再設定アクション
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.

分類	説明	再設定アクション
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.9.0 変更ログ

### 6.9.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-08-30	リリースノートの更新	タイミングシーケンスの不一致の問題の修正が追加されました。
2023-08-21	リリースノートの更新	Hadoop 3.3.3 に関する既知の問題が追加されました。
2023-07-26	更新	新しい OS リリースラベル 2.0.20230612.0 と 2.0.20230628.0 。
2022-12-13	リリースノートが更新されました	でランタイムに関する機能と既知の問題を追加 SageMaker
2022-11-29	リリースノートとドキュメントが更新されました	Apache Spark 用の Amazon Redshift の統合の機能が追加されました
2022-11-23	リリースノートが更新されました	Log4j エントリが削除されました
2022-11-18	デプロイ完了	Amazon EMR 6.9 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>

日付	イベント	説明
2022-11-18	ドキュメント公開	Amazon EMR 6.9 リリース ノートの初版が公開されました
2022-11-14	初回リリース	Amazon EMR 6.9 が限定商用リージョンにデプロイされました

## Amazon EMR リリース 6.8.1

### 6.8.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)、[Ju](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7



	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
Scala	2.12.15	2.12.15	2.12.15	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.15.1	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.12	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.3	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.2
Hudi	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.273	0.273	0.272	0.267
Spark	3.3.0	3.3.0	3.2.1	3.2.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.9.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	388	378	367
Zeppelin	0.10.1	0.10.1	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.7	3.5.7

## 6.8.1 リリースノート

次のリリースノートには、Amazon EMR リリース 6.8.1 に関する情報が含まれています。6.8.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.8.1 変更ログ](#)」を参照してください。

### 変更点、機能強化、解決した問題

- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。このアプローチでは、マネージドスケールリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。

Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために `yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値が `false` に設定されています。

この修正は YARN-9608 によって発生した問題を解決するものですが、マネージドスケールリングが有効になっているクラスターでのシャッフルデータ損失が原因で Hive ジョブが失敗する可能性があります。このリリースでは、Hive ワークロードにも `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` を設定することで、このリスクを軽減しました。この設定は、Amazon EMR リリース 6.11.0 以降でのみ使用できます。

- インスタンスグループが設定されているクラスター内のプライマリノードのフェイルオーバー後、メトリクスコレクターはコントロールプレーンにメトリクスを送信しません。
- このリリースでは、メトリクスコレクターのエンドポイントへの HTTP リクエストが失敗しても再試行する必要がなくなりました。
- このリリースには、高可用性クラスターが再起動後に障害状態から回復できるようにする変更が含まれています。
- このリリースでは、ユーザーが作成した大きな UID がオーバーフロー例外を引き起こしていた問題が修正されています。
- このリリースでは、Amazon EMR 再構成プロセスのタイムアウトの問題が修正されています。
- このリリースでは、再構成に失敗したために他の無関係なプロセスが中断される可能性があるという問題が防止されます。
- このリリースにはセキュリティ修正が含まれています。
- このリリースでは、Amazon EMR で Spark でワークロードを実行しているクラスターが `contains`、`startsWith`、`endsWith`、および `like` でメッセージが表示されずに間違った結果を受け取る可能性があるという問題が修正されています。この問題は、Amazon EMR Hive3 メタストアサーバー (HMS) のメタデータを含むパーティション化されたフィールドで式を使用する場合に発生します。
- Amazon EMR 6.6.0 から 6.9.x では、動的パーティションと ORDER BY 句または SORT BY 句を使用した INSERT クエリには常に 2 つのリデューサーがあります。この問題は、OSS が変更された [HIVE-20703](#) が原因です。これにより、動的ソートパーティションの最適化がコストベースの決定下に置かれます。ワークロードで動的パーティションのソートが不要な場合は、`hive.optimize.sort.dynamic.partition.threshold` プロパティを `-1` に設定して新機能を無効にし、リデューサーの数を正しく計算することをお勧めします。この問題は、[HIVE-22269](#) の一部として OSS Hive で修正され、Amazon EMR 6.10.0 で修正されています。
- HDFS をスクラッチディレクトリとして使用し、小さいファイルのマージを有効にしていて、テーブルに静的なパーティションパスが含まれている場合、Hive でデータが失われる可能性があります。

- このリリースでは、ETL ジョブの終了時に小さいファイルのマージ (デフォルトでは無効) が有効になっている場合の Hive のパフォーマンスの問題が修正されています。
- このリリースでは、ユーザー定義関数 (UDF) がない場合の Glue 側のスロットリングの問題が修正されています。
- このリリースでは、YARN が廃止された場合に、ログプッシャーが S3 にコンテナログをプッシュする前に、ノードログ集約サービスによってコンテナログが削除される問題が修正されています。
- このリリースでは、HBase の永続的なストアファイル追跡による圧縮/アーカイブファイルの処理が修正されています。
- このリリースでは、spark-defaults.conf の spark.yarn.heterogeneousExecutors.enabled 設定にデフォルトの true 値を設定したときに Spark のパフォーマンスに影響する問題が修正されています。
- このリリースでは、低減タスクがシャッフルデータの読み取りに失敗する問題が修正されています。この問題により、メモリ破損エラーで Hive のクエリが失敗していました。
- このリリースでは、ノード交換中に HDFS NameNode (NN) サービスがセーフモードで停止した場合にノードプロビジョナーが失敗する問題が修正されています。
- このリリースでは、Presto または Trino を実行する EMR クラスターのクラスタースケールリングワークフローに新しい再試行メカニズムが追加されています。この改善により、サイズ変更操作が 1 回失敗したためにクラスターのサイズ変更が無期限に停止するリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。
- このリリースでは、クラスターのスケールダウンロジックが改善され、クラスターの HDFS レプリケーション係数設定を下回るコアノードのスケールダウンが試みられることがなくなりました。これはデータの冗長性要件に合致し、スケールリング操作が停止する可能性が低くなります。
- ログ管理デーモンがアップグレードされ、ローカルインスタンスストレージ上のオープンファイルハンドルでアクティブに使用されているすべてのログと、関連するプロセスが識別されるようになりました。このアップグレードにより、ログが Amazon S3 にアーカイブされた後に Amazon EMR がファイルを適切に削除し、ストレージスペースを再利用できるようになります。
- このリリースには、ローカルクラスターファイルシステム内の空で未使用のステップディレクトリを削除するログ管理デーモンの機能強化が含まれています。空のディレクトリが多すぎると、Amazon EMR デーモンのパフォーマンスが低下し、ディスクが過剰に使用される可能性があります。
- このリリースでは、複数のプライマリノードがあるクラスターからプライマリノードの 1 つを複製してエッジノードを作成する際に発生する可能性のある問題が修正されています。複製されたエッジノードは、スケールダウン操作で遅延を引き起こしたり、プライマリノードのメモリ使用率が高くなったりする可能性があります。EMR クラスターと通信するエッジノードを作成する方

法の詳細については、の `aws-samples` リポジトリの「[エッジノードの作成](#)」を参照してください  
GitHub。

- このリリースでは、Amazon EMR が再起動後に Amazon EBS ボリュームをインスタンスに再マウントするために使用する自動化プロセスが改善されました。
- このリリースでは、Amazon EMR が Amazon に発行する Hadoop メトリクスに断続的なギャップが生じる問題が修正されています CloudWatch。
- このリリースでは、クラスターのノードの除外リストを含む YARN 構成ファイルの更新が、ディスクの過剰使用により中断される EMR クラスターの問題が修正されています。更新が不完全だと、今後のクラスターのスケールダウン操作が妨げられます。このリリースでは、クラスターが正常に動作し、スケールダウン操作が期待どおりに機能することが保証されます。
- このリリースでは、クラスター上のログ管理デーモンが改善され、EMR クラスター内の追加のログフォルダを監視できるようになりました。この改善により、ディスクの過剰使用シナリオが最小限に抑えられます。
- このリリースでは、クラスター上のログ管理デーモンが停止すると、自動的に再起動されます。この改善により、ディスクの過剰使用が原因でノードが異常に見えるリスクが軽減されます。
- このリリースでは、クラスターのスケールダウン中に Amazon S3 にログをアーカイブするためのサポートが追加されています。以前は、Amazon S3 にログファイルをアーカイブできるのはクラスター終了時のみでした。この新機能により、クラスターで生成されたログファイルは、ノードが終了した後も Amazon S3 に残ります。詳細については、「[クラスターのログ記録とデバッグを設定する](#)」を参照してください。
- このリリースでは、ブートストラップアクションの Amazon S3 URI がポート番号で終わっている場合 (例: `a.b.c.d:4345`) に発生する問題が修正されています。Amazon EMR は、このような URI を誤って解析していたため、関連するブートストラップアクションはすべて失敗していました。
- このリリースでは、Apache BigTop と EC2 クラスターの Amazon EMR の起動シーケンス間のタイミングシーケンスの不一致の問題が修正されています。このタイミングシーケンスの不一致は、システムが 2 つ以上の操作を適切なシーケンスで実行するのではなく、同時に実行しようとした場合に発生します。その結果、特定のクラスター構成でインスタンスの起動がタイムアウトし、クラスターの起動時間が遅くなりました。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			GovCloud (米国西部)、AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、



OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、



OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			AWS GovCloud (米国東部)、中国(北京)、中国(寧夏)
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

### 6.8.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-notebook-env</code>	1.7.0	Jupyter エンタープライズゲートウェイを含む <code>emr</code> ノートブック用 Conda env
<code>emr-s3-dist-cp</code>	2.22.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	2.1.0	EMR S3Select コネクタ

コンポーネント	Version	説明
emrfs	2.53.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.15.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.15.1	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-8.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-8.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-8.1	HDFS コマンドラインクライアントとライブラリ



コンポーネント	Version	説明
hadoop-hdfs-namenode	3.2.1-amzn-8.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-8.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httfs-server	3.2.1-amzn-8.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-8.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-8.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-8.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-8.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-8.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.12-amzn-0.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。

コンポーネント	Version	説明
hbase-region-server	2.4.12-amzn-0.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.12-amzn-0.1	HBase コマンドラインクライアント。
hbase-rest-server	2.4.12-amzn-0.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.12-amzn-0.1	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.12-amzn-0.1	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-1.1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-1.1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-1.1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-1.1	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-1.1	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	3.1.3-amzn-1.1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-1.1	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.11.1-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.11.1-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.11.1-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.11.1-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	0.14.0-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.7.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.2	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.273.3-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.273.3-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.273.3-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	388-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	388-amzn-0	いくつかのクエリを実行するサービス。
trino-client	388-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.3.0-amzn-0.1	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	3.3.0-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.3.0-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.3.0-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	22.06.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.9.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

### 6.8.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-6.8.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ

分類	説明	再設定アクション
		er, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.



分類	説明	再設定アクション
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.

分類	説明	再設定アクション
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.

分類	説明	再設定アクション
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)

分類	説明	再設定アクション
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の musql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-log4j	Ranger KMS の kms-log4j .properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.8.1 変更ログ

### 6.8.1 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-08-30	リリースノートの更新	リリースノートにコントロールプレーン関連の修正をいくつか追加しました
2023-08-21	ドキュメント公開	Amazon EMR 6.8.1 リリースノートの初版が公開されました
2023-08-16	デプロイ完了	Amazon EMR 6.8.1 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-08-04	初回リリース	Amazon EMR 6.8.1 が限定商用リージョンに初めてデプロイされました

## Amazon EMR リリース 6.8.0

### 6.8.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)、[Ju](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.31
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.15	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.14.2	1.14.2	1.14.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.4	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.2	3.1.2
Hudi	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1	0.12.0
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.273	0.272	0.267	0.261
Spark	3.3.0	3.2.1	3.2.0	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	378	367	360
Zeppelin	0.10.1	0.10.0	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.7	3.5.7	3.5.7

## 6.8.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.8.0 に関する情報が含まれています。6.7.0 からの変更が含まれています。

### 新機能

- Amazon EMR ステップ機能が Apache Livy エンドポイントと JDBC/ODBC クライアントをサポートするようになりました。詳細については、「[Configure runtime roles for Amazon EMR steps](#)」を参照してください。

- Amazon EMR リリース 6.8.0 には Apache HBase リリース 2.4.12 が付属しています。この HBase リリースでは、HBase テーブルのアーカイブと削除の両方を行うことができます。Amazon S3 アーカイブプロセスでは、すべてのテーブルファイルの名前がアーカイブディレクトリに変更されます。このプロセスにはコストがかかり、時間を要する場合があります。これで、アーカイブプロセスをスキップして、大きなテーブルをすばやく削除することができます。詳細については、「[HBase シェルを使用する](#)」を参照してください。

## 既知の問題

- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。Amazon EMR 6.8.0 および 6.9.0 の場合、この方法では、マネージドスケーリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。


[Amazon EMR 6.10.0](#) では、`yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値を `false` に設定することでこの問題を回避できます。Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために設定がデフォルトで `false` にセットされています。

## 変更、拡張、解決した問題

- Amazon EMR リリース 6.5.0、6.6.0、または 6.7.0 が Apache Spark シェルを介して Apache Phoenix テーブルを読み取ると、Amazon EMR は `NoSuchMethodError` を生成しました。Amazon EMR リリース 6.8.0 では、この問題が修正されています。
- Amazon EMR リリース 6.8.0 には [Apache Hudi](#) 0.11.1 が付属していますが、Amazon EMR 6.8.0 クラスターは Hudi 0.12.0 のオープンソース `hudi-spark3.3-bundle_2.12` と互換性があります。
- Amazon EMR リリース 6.8.0 には、Apache Spark 3.3.0 が付属しています。この Spark リリースでは、Apache Log4j 2 と `log4j2.properties` ファイルを使用して Spark プロセス内の Log4j を設定します。クラスターで Spark を使用するか、カスタム設定パラメータを使用して EMR クラスターを作成し、Amazon EMR リリース 6.8.0 にアップグレードする場合は、Apache Log4j 2 の新しい `spark-log4j2` 設定分類とキー形式に移行する必要があります。詳細については、「[Apache Log4j 1.x から Log4j 2.x への移行](#)」を参照してください。



- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

 Note

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または [ListReleaseLabels](#) API あるいは [list-release-labels](#) CLI アクションを使用するとき Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			ン)、中東 (アラブ首長国連邦)、カナダ (中部)
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202307.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 119.1	4.14.301	2023 年 2 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022004.0	4.14.294	2022 年 11 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 912.1	4.14.291	2022 年 9 月 6 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

## 既知の問題

- Spark on Amazon EMR バージョン 6.6.0、6.7.0、6.8.0 で DynamoDB コネクタを使用すると、テーブルから何を読み込んでも空の結果が返されます。この状況は、入力分割が空でないデータを参照している場合でも変わりません。これは、Spark 3.2.0 が `spark.hadoopRDD.ignoreEmptySplits` をデフォルトで `true` に設定しているためです。回避策として、明示的に `spark.hadoopRDD.ignoreEmptySplits` を `false` に設定します。Amazon EMR リリース 6.9.0 では、この問題が修正されています。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、Log4j2 プロパティのファイル名パターン設定が正しくないため、SecretAgent および RecordServer サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

## 6.8.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.2	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.2.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.7.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.22.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.1.0	EMR S3Select コネクタ
emrfs	2.53.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.15.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.15.1	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-8	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-8	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-8	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-8	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-8	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-8	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-8	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-8	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-8	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	3.2.1-amzn-8	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-8	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.12-amzn-0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.12-amzn-0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.12-amzn-0	HBase コマンドラインクライアント。
hbase-rest-server	2.4.12-amzn-0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.12-amzn-0	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.12-amzn-0	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.3-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	3.1.3-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-1	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-1	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.11.1-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.11.1-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.11.1-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.11.1-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション

コンポーネント	Version	説明
iceberg	0.14.0-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.9.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.7.0	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ



コンポーネント	Version	説明
phoenix-connectors	5.1.2	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.273.3-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.273.3-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.273.3-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	388-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	388-amzn-0	いくつかのクエリを実行するサービス。
trino-client	388-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。

コンポーネント	Version	説明
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.3.0-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.3.0-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.3.0-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.3.0-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	22.06.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.9.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.10.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.10	ZooKeeper コマンドラインクライアント。

## 6.8.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

### emr-6.8.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

分類	説明	再設定アクション
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

分類	説明	再設定アクション
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

分類	説明	再設定アクション
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.

分類	説明	再設定アクション
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.

分類	説明	再設定アクション
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.

分類	説明	再設定アクション
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.

分類	説明	再設定アクション
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j2	Spark の log4j2.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.



分類	説明	再設定アクション
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 6.8.0 変更ログ

### 6.8.0 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-08-21	更新	Hadoop 3.3.3 に関する既知の問題が追加されました。
2023-07-26	更新	新しい OS リリースラベル 2.0.20230612.0 と 2.0.20230628.0 。
2022-09-06	デプロイ完了	Amazon EMR 6.8 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2022-09-06	初版発行	Amazon EMR 6.8 リリースノートの初版が公開されました
2022-08-31	初回リリース	Amazon EMR 6.8 が限定商用リージョンにリリースされました

## Amazon EMR リリース 6.7.0

### 6.7.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)、[Ju](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
AWS SDK for Java	1.12.170	1.12.170	1.12.31	1.12.31
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.15	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.0	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.4.4

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
HCatalog	3.1.3	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.2	3.1.2	3.1.2
Hudi	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	0.13.1-amzn-0	0.13.1	0.12.0	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.272	0.267	0.261	0.254.1
Spark	3.2.1	3.2.0	3.1.2	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
Trino (PrestoSQL)	378	367	360	359
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.5.7

## 6.7.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.7.0 に関する情報が含まれています。6.6.0 からの変更が含まれています。

初回リリース日: 2022 年 7 月 15 日

### 新機能

- Amazon EMR は Apache Spark 3.2.1、Apache Hive 3.1.3、HUDI 0.11、PrestoDB 0.272、Trino 0.378 をサポートするようになりました。
- EC2 クラスター上の Amazon EMR 用の EMR ステップ (Spark、Hive) による IAM ロールおよび Lake Formation ベースのアクセスコントロールをサポートします。
- Apache Ranger 対応クラスターでの Apache Spark データ定義ステートメントをサポートします。これには、Apache Ranger 対応クラスターで Apache Hive メタデータを読み書きする Trino アプリケーションのサポートが含まれるようになりました。詳細については、「[Enable federated governance using Trino and Apache Ranger on Amazon EMR](#)」を参照してください。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			( 米国東部 )、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 207.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023119.1	4.14.301	2023 年 2 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 004.0	4.14.294	2022 年 11 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 912.1	4.14.291	2022 年 10 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 719.0	4.14.287	2022 年 8 月 10 日	us-west-1 , eu-west-3 , eu-north-1 , ap-south-1 , me-south-1

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 606.1	4.14.281	2022 年 7 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

## 既知の問題

- Amazon EMR リリース 6.5.0、6.6.0、または 6.7.0 が Apache Spark シェルを介して Apache Phoenix テーブルを読み取ると、Amazon EMR が誤った `Hbase.compat.version` を使用しているため、`NoSuchMethodError` が発生します。Amazon EMR リリース 6.8.0 では、この問題が修正されています。
- Spark on Amazon EMR バージョン 6.6.0、6.7.0、6.8.0 で DynamoDB コネクタを使用すると、テーブルから何を読み込んでも空の結果が返されます。この状況は、入力分割が空でないデータを参照している場合でも変わりません。これは、Spark 3.2.0 が `spark.hadoopRDD.ignoreEmptySplits` をデフォルトで `true` に設定しているためです。回避策として、明示的に `spark.hadoopRDD.ignoreEmptySplits` を `false` に設定します。Amazon EMR リリース 6.9.0 では、この問題が修正されています。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、`Log4j2` プロパティのファイル名パターン設定が正しくないため、`SecretAgent` および `RecordServer` サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- GetClusterSessionCredentials API は、Amazon EMR 6.7 以前で実行されるクラスターではサポートされていません。

## 6.7.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK

コンポーネント	Version	説明
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.6.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.22.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.1.0	EMR S3Select コネクタ
emrfs	2.52.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.14.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.14.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-7	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-7	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-7	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-7	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-7	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-7	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-7	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。



コンポーネント	Version	説明
hadoop-mapred	3.2.1-amzn-7	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-7	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-7	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-7	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.4-amzn-3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.4-amzn-3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.4-amzn-3	HBase コマンドラインクライアント。
hbase-rest-server	2.4.4-amzn-3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.4-amzn-3	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.4-amzn-3	Apache HBase クラスター用の修復ツール。

コンポーネント	Version	説明
hcatalog-client	3.1.3-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.3-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.3-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.3-amzn-0	Hive コマンドラインクライアント。
hive-hbase	3.1.3-amzn-0	Hive-hbase クライアント。
hive-metastore-server	3.1.3-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.3-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.11.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.11.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.11.0-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hudi-spark	0.11.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	0.13.1-amzn-0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.8.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.0.194	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-connectors	5.1.2	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.272-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.272-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.272-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	378-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	378-amzn-0	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
trino-client	378-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.2.1-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.2.1-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.2.1-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.2.1-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	22.02.0-amzn-1	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.7	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.7	ZooKeeper コマンドラインクライアント。

## 6.7.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-6.7.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.



分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.

分類	説明	再設定アクション
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.

分類	説明	再設定アクション
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)

分類	説明	再設定アクション
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の musql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-log4j	Ranger KMS の kms-log4j .properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 6.6.0

### 6.6.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)、[Ju](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
AWS SDK for Java	1.12.170	1.12.31	1.12.31	1.11.977
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.0	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
Hudi	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	0.13.1	0.12.0	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.267	0.261	0.254.1	0.245.1
Spark	3.2.0	3.1.2	3.1.2	3.1.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	367	360	359	350
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.4.14

## 6.6.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.6.0 に関する情報が含まれています。6.5.0 からの変更が含まれています。

初回リリース日: 2022 年 5 月 9 日

ドキュメント更新日: 2022 年 6 月 15 日

### 新機能

- Amazon EMR 6.6 は、Apache Spark 3.2、Apache Spark RAPIDS 22.02、CUDA 11、Apache Hudi 0.10.1、Apache Iceberg 0.13、Trino 0.367、PrestoDB 0.267 をサポートするようになりました。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、ア

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)</p>



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 207.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 119.1	4.14.301	2023 年 2 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 004.0	4.14.294	2022 年 11 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 912.1	4.14.291	2022 年 10 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2022 805.0	4.14.287	2022 年 8 月 30 日	us-west-1

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 719.0	4.14.287	2022 年 8 月 10 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 426.0	4.14.281	2022 年 6 月 10 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 406.1	4.14.275	2022 年 5 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

- Amazon EMR 6.6 以降では、Log4j 1.x と Log4j 2.x を使用するアプリケーションは、それぞれ Log4j 1.2.17 (またはそれ以上) と Log4j 2.17.1 (またはそれ以上) を使用するようにアップグレードされます。CVE の問題を軽減するために提供されている [ブートストラップアクション](#) を使用する必要はなくなります。

- [マネージドスケールリング] Spark シャッフルデータマネージドスケールリング最適化 - Amazon EMR バージョン 5.34.0 以降、および EMR バージョン 6.4.0 以降では、マネージドスケールリングは Spark シャッフルデータ (Spark が特定の操作を実行するためにパーティション間で再分配するデータ) 対応になりました。シャッフル操作の詳細については、「Amazon EMR 管理ガイド」および「[Spark のプログラミングガイド](#)」の「[Amazon EMR での EMR マネージドスケールリングの使用](#)」を参照してください。
- Amazon EMR 5.32.0 および 6.5.0 以降、Apache Spark の動的エグゼキューターサイズ設定はデフォルトで有効になっています。この機能をオンまたはオフにするには、`spark.yarn.heterogeneousExecutors.enabled` 設定パラメータを使用できます。

## 変更、拡張、解決した問題

- Amazon EMR は、EMR のデフォルト AMI オプションを使用し、Apache Hadoop、Apache Spark、Apache Hive などの一般的なアプリケーションのみをインストールするクラスターの場合、クラスターの起動時間を平均で最大 80 秒短縮します。

## 既知の問題

- Amazon EMR リリース 6.5.0、6.6.0、または 6.7.0 が Apache Spark シェルを介して Apache Phoenix テーブルを読み取ると、Amazon EMR が誤った `Hbase.compat.version` を使用しているため、`NoSuchMethodError` が発生します。Amazon EMR リリース 6.8.0 では、この問題が修正されています。
- Spark on Amazon EMR バージョン 6.6.0、6.7.0、6.8.0 で DynamoDB コネクタを使用すると、テーブルから何を読み込んでも空の結果が返されます。この状況は、入力分割が空でないデータを参照している場合でも変わりません。これは、Spark 3.2.0 が `spark.hadoopRDD.ignoreEmptySplits` をデフォルトで `true` に設定しているためです。回避策として、明示的に `spark.hadoopRDD.ignoreEmptySplits` を `false` に設定します。Amazon EMR リリース 6.9.0 では、この問題が修正されています。
- Trino の長時間稼働クラスターでは、Amazon EMR 6.6.0 は Trino `jvm.config` のガベージコレクションログ記録パラメータを有効にして、ガベージコレクションログからより深い知見を得ることができます。この変更により、`launcher.log (/var/log/trino/launcher.log)` ファイルに多くのガベージコレクションログが追加されます。Amazon EMR 6.6.0 で Trino クラスターを実行している場合、追加されたログが原因で、クラスターを数日間実行した後にノードのディスク容量が不足することがあります。



この問題の回避策は、Amazon EMR 6.6.0 のクラスターを作成またはクローン作成する際に、以下のスクリプトをブートストラップアクションとして実行し、`jvm.config` のガベージコレクションログ記録パラメータを無効にすることです。

```
#!/bin/bash
set -ex
PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、`Log4j2` プロパティのファイル名パターン設定が正しくないため、`SecretAgent` および `RecordServer` サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

## 6.6.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.2.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.5.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.20.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.1.0	EMR S3Select コネクタ
emrfs	2.50.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.14.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.14.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-6	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-6	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-6	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	3.2.1-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.4-amzn-2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.4-amzn-2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.4-amzn-2	HBase コマンドラインクライアント。
hbase-rest-server	2.4.4-amzn-2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.4-amzn-2	HBase に Thrift エンドポイントを提供するサービス。
hbase-operator-tools	2.4.4-amzn-2	Apache HBase クラスター用の修復ツール。
hcatalog-client	3.1.2-amzn-7	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-7	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	3.1.2-amzn-7	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-7	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-7	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-7	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-7	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.10.1-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.10.1-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.10.1-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.10.1-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション

コンポーネント	Version	説明
iceberg	0.13.1	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.8.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	11.0.194	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-connectors	5.1.2	Spark-3 用 Apache Phoenix-Connectors
phoenix-query-server	5.1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.267-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.267-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.267-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	367-amzn-0	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	367-amzn-0	いくつかのクエリを実行するサービス。
trino-client	367-amzn-0	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。



コンポーネント	Version	説明
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.2.0-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.2.0-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.2.0-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.2.0-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	22.02.0-amzn-0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.10.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.7	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.7	ZooKeeper コマンドラインクライアント。

### 6.6.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-6.6.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

分類	説明	再設定アクション
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

分類	説明	再設定アクション
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

分類	説明	再設定アクション
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.

分類	説明	再設定アクション
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.



分類	説明	再設定アクション
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.

分類	説明	再設定アクション
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino の iceberg.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.

分類	説明	再設定アクション
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.

分類	説明	再設定アクション
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 6.5.0

### 6.5.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)、[Ju](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)



## アプリケーションバージョン情報

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
AWS SDK for Java	1.12.31	1.12.31	1.11.977	1.11.977
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.0	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.2.6	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0
Hue	4.9.0	4.9.0	4.9.0	4.9.0
Iceberg	0.12.0	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.261	0.254.1	0.245.1	0.245.1
Spark	3.1.2	3.1.2	3.1.1	3.1.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	360	359	350	350
Zeppelin	0.10.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.4.14	3.4.14

## 6.5.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.5.0 に関する情報が含まれています。6.4.0 からの変更が含まれています。

初回リリース日: 2022 年 1 月 20 日

更新リリース日: 2022 年 3 月 21 日

### 新機能

- [マネージドスケールリング] Spark シャッフルデータマネージドスケールリング最適化 - Amazon EMR バージョン 5.34.0 以降、および EMR バージョン 6.4.0 以降では、マネージドスケールリングは Spark シャッフルデータ (Spark が特定の操作を実行するためにパーティション間で再分配する

データ) 対応になりました。シャッフル操作の詳細については、「Amazon EMR 管理ガイド」および「[Spark のプログラミングガイド](#)」の「[Amazon EMR での EMR マネージドスケーリングの使用](#)」を参照してください。

- Amazon EMR 5.32.0 および 6.5.0 以降、Apache Spark の動的エグゼキューターサイズ設定はデフォルトで有効になっています。この機能をオンまたはオフにするには、`spark.yarn.heterogeneousExecutors.enabled` 設定パラメータを使用できます。
- 巨大な分析データセット用の Apache Iceberg オープンテーブル形式のサポート。
- ranger-trino-plugin 2.0.1-amzn-1 のサポート
- toree 0.5.0 のサポート

### 変更、拡張、解決した問題

- Amazon EMR 6.5 リリースバージョンは Apache Iceberg 0.12.0 をサポートするようになり、Apache Spark 用 Amazon EMR ランタイム、Presto 用 Amazon EMR ランタイム、Apache Hive 用 Amazon EMR ランタイムによりランタイムが改善されています。
- [Apache Iceberg](#) は、Amazon S3 の大規模データセット用のオープンテーブル形式であり、大きなテーブルに対する高速のクエリパフォーマンス、アトミックコミット、同時書き込み、および SQL 互換テーブルの進化を提供します。EMR 6.5 では、Apache Spark 3.1.2 を Iceberg テーブル形式で使用できます。
- Apache Hudi 0.9 では Spark SQL DDL と DML のサポートが追加されています。これにより、SQL ステートメントだけを使用して Hudi テーブルを作成し、アップサートできるようになりました。Apache Hudi 0.9 では、クエリ側とライター側のパフォーマンスも向上しています。
- Apache Hive 用 Amazon EMR ランタイムは、ステージング操作中の名前変更操作を削除することで Amazon S3 での Apache Hive のパフォーマンスを向上させ、テーブルの修復に使用されるメタストアチェック (MSCK) コマンドのパフォーマンスを向上させます。

### 既知の問題

- Amazon EMR リリース 6.5.0、6.6.0、または 6.7.0 が Apache Spark シエルを介して Apache Phoenix テーブルを読み取ると、Amazon EMR が誤った `Hbase.compat.version` を使用しているため、`NoSuchMethodError` が発生します。Amazon EMR リリース 6.8.0 では、この問題が修正されています。
- 高可用性 (HA) の Hbase バンドルクラスターは、デフォルトのボリュームサイズとインスタンスタイプではプロビジョニングに失敗します。この問題の回避策は、ルートボリュームのサイズを増やすことです。

- Apache Oozie で Spark アクションを使用するには、以下の設定を Oozie workflow.xml ファイルに追加する必要があります。そうしないと、Hadoop や EMRFS などのいくつかの重要なライブラリが、Oozie が起動する Spark エグゼキューターのクラスパスから失われてしまいます。

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、s3://bucket/table/p=a は s3://bucket/table/p=a b のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の spark.sql.sources.fastS3PartitionDiscovery.enabled 設定を false にセットします。

## 6.5.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオー

プンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.4.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.19.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.1.0	EMR S3Select コネクタ
emrfs	2.48.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.14.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.14.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。

コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-5	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-5	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-5	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-5	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-5	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-5	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	3.2.1-amzn-5	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-5	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-5	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-5	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-5	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.4-amzn-1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.4.4-amzn-1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.4-amzn-1	HBase コマンドラインクライアント。
hbase-rest-server	2.4.4-amzn-1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.4-amzn-1	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	3.1.2-amzn-6	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-6	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.2-amzn-6	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-6	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-6	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-6	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-6	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.9.0-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.9.0-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.9.0-amzn-1	Trino を Hudi で実行するためのバンドルライブラリ。



コンポーネント	Version	説明
hudi-spark	0.9.0-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.9.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
iceberg	0.12.0	Apache Iceberg は、膨大な分析データセット用のオープンテーブル形式です。
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.8.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.261-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.261-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.261-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	360	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	360	いくつかのクエリを実行するサービス。
trino-client	360	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。

コンポーネント	Version	説明
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.1.2-amzn-1	Spark コマンドラインクライアント。
spark-history-server	3.1.2-amzn-1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.1.2-amzn-1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.1.2-amzn-1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	0.4.1	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.10.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.7	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.7	ZooKeeper コマンドラインクライアント。

## 6.5.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

### emr-6.5.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.

分類	説明	再設定アクション
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.



分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
iceberg-defaults	Iceberg の iceberg-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.

分類	説明	再設定アクション
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.

分類	説明	再設定アクション
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)

分類	説明	再設定アクション
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の musql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.



分類	説明	再設定アクション
ranger-kms-log4j	Ranger KMS の kms-log4j .properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 6.4.0

### 6.4.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
AWS SDK for Java	1.12.31	1.11.977	1.11.977	1.11.880
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.2.6	2.2.6	2.2.6-amzn-0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
Hudi	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	5.1.2	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.254.1	0.245.1	0.245.1	0.238.3
Spark	3.1.2	3.1.1	3.1.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	359	350	350	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.4.14	3.4.14	3.4.14

## 6.4.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.4.0 に関する情報が含まれています。6.3.0 からの変更が含まれています。

初回リリース日: 2021 年 9 月 20 日

更新リリース日: 2022 年 3 月 21 日

サポートされているアプリケーション

- AWS SDK for Java バージョン 1.12.31
- CloudWatch シンクバージョン 2.2.0
- DynamoDB Connector バージョン 4.16.0
- EMRFS バージョン 2.47.0
- Amazon EMR Goodies バージョン 3.2.0
- Amazon EMR Kinesis Connector バージョン 3.5.0
- Amazon EMR Record Server バージョン 2.1.0
- Amazon EMR Scripts バージョン 2.5.0
- Flink バージョン 1.13.1
- Ganglia バージョン 3.7.2
- AWS Glue Hive メタストアクライアントバージョン 3.3.0
- Hadoop バージョン 3.2.1-amzn-4
- HBase バージョン 2.4.4-amzn-0
- HB ase-operator-tools 1.1.0
- HCatalog バージョン 3.1.2-amzn-5
- Hive バージョン 3.1.2-amzn-5
- Hudi バージョン 0.8.0-amzn-0
- Hue バージョン 4.9.0
- Java JDK バージョン Corretto-8.302.08.1 (ビルド 1.8.0\_302-b08)
- JupyterHub バージョン 1.4.1
- Livy バージョン 0.7.1-incubating

- MXNet バージョン 1.8.0
- Oozie バージョン 5.2.1
- Phoenix バージョン 5.1.2
- Pig バージョン 0.17.0
- Presto バージョン 0.254.1-amzn-0
- Trino バージョン 359
- Apache Ranger KMS (マルチマスター透過的暗号化) バージョン 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker Spark SDK バージョン 1.4.1
- Scala バージョン 2.12.10 (OpenJDK 64 ビットサーバー VM、Java 1.8.0\_282)
- Spark バージョン 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop バージョン 1.4.7
- TensorFlow バージョン 2.4.1
- tez バージョン 0.9.2
- Zeppelin バージョン 0.9.0
- Zookeeper バージョン 3.5.7
- コネクタおよびドライバー: DynamoDB Connector 4.16.0

## 新機能

- [マネージドスケールリング] Spark シャッフルデータマネージドスケールリング最適化 - Amazon EMR バージョン 5.34.0 以降、および EMR バージョン 6.4.0 以降では、マネージドスケールリングは Spark シャッフルデータ (Spark が特定の操作を実行するためにパーティション間で再分配するデータ) 対応になりました。シャッフル操作の詳細については、「[Amazon EMR 管理ガイド](#)」および「[Spark のプログラミングガイド](#)」の「[Amazon EMR での EMR マネージドスケールリングの使用](#)」を参照してください。
- Apache Ranger 対応の Amazon EMR クラスターでは、Apache Spark SQL を使用して、INSERT INTO、INSERT OVERWRITE、および ALTER TABLE で Apache Hive メタストアテーブルでデータの挿入または更新を実行できます。Spark SQL で ALTER TABLE を使用する場合、パーティションの場所はテーブルの場所の子ディレクトリである必要があります。Amazon EMR は現在、

パーティションの場所がテーブルの場所と異なるパーティションへのデータの挿入をサポートしていません。

- PrestoSQL は [Trino に名称変更](#)されています。
- Hive: LIMIT 句を使用した簡単な SELECT クエリの実行は、LIMIT 句に記述されているレコード数がフェッチされたらすぐにクエリの実行を停止すると、高速化されます。簡単な SELECT クエリは、GROUP BY/ORDER BY 句のないクエリ、またはリデュースーステージのないクエリです。例えば SELECT \* from <TABLE> WHERE <Condition> LIMIT <Number> です。

## Hudi の同時実行制御

- Hudi はオプティミスティック同時実行制御 (OCC) をサポートするようになりました。この OCC を UPSERT や INSERT などの書き込みオペレーションで利用して、複数のライターから同じ Hudi テーブルへの変更を許可できます。これはファイルレベルの OCC であるため、2 つのコミット (またはライター) は、その変更が競合しなければ、同じテーブルに書き込むことができます。詳細については、「[Hudi Concurrency Control](#)」を参照してください。
- Amazon EMR クラスターには Zookeeper がインストールされており、OCC のロックプロバイダーとして利用できます。この機能を使いやすくするために、Amazon EMR クラスターには次のプロパティが事前設定されています。

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

OCC を有効にするには、Hudi ジョブオプションを使用して、または Amazon EMR 設定 API を使用してクラスターレベルで、次のプロパティを設定する必要があります。

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table Name is a good option)
```

## Hudi モニタリング: Hudi メトリクスを報告するための Amazon CloudWatch 統合

- Amazon EMR は、Amazon への Hudi メトリクスの発行をサポートしています CloudWatch。これを有効にするには、次の必要な設定を行います。

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- 変更できるオプションの Hudi 設定を以下に示します。

設定	説明	値
hoodie.metrics.cloudwatch.report.period.seconds	Amazon にメトリクスを報告する頻度 (秒単位) CloudWatch	デフォルト値は 60 秒で、Amazon が提供するデフォルトの 1 分間の解像度でも問題ありません。CloudWatch
hoodie.metrics.cloudwatch.metric.prefix	各メトリクス名に追加するプレフィックス	デフォルト値は空です (プレフィックスなし)
hoodie.metrics.cloudwatch.namespaces	メトリクスが公開される Amazon CloudWatch 名前空間	デフォルト値は Hudi です
hoodie.metrics.cloudwatch.maxDatumsPerRequest	Amazon への 1 つのリクエストに含めるデータムの最大数 CloudWatch	デフォルト値は 20 で、Amazon の CloudWatch デフォルトと同じです。

## Amazon EMR Hudi の設定のサポートと改善

- EMR 設定 API と再設定機能を利用して、クラスターレベルで Hudi 設定を構成できるようになりました。Spark、Hive などの他のアプリケーションに似た `/etc/hudi/conf/hudi-defaults.conf` を使用して、新しいファイルベースの設定のサポートが導入されました。EMR では、ユーザーエクスペリエンスを向上させるためにデフォルト設定はほとんどありません。

— `hoodie.datasource.hive_sync.jdbcurl` はクラスター Hive サーバ URL に設定されるため、指定する必要がなくなりました。これは、Spark クラスターモードでジョブを実行する場合に特に便利です。この場合、以前は Amazon EMR マスター IP を指定する必要がありました。

— HBase 固有の設定。Hbase のインデックスを Hudi で使用する場合に役立ちます。



- 同時実行制御で説明されているように、Zookeeper はプロバイダー固有の設定をロックします。これにより、オプティミスティック同時実行制御 (OCC) の使用が容易になります。
- 渡す必要がある設定の数を減らし、可能な場合は自動的に推測するために、追加の変更が導入されました。
  - `partitionBy` キーワードは、パーティション列を指定するために使用できます。
  - Hive Sync を有効にすると、`HIVE_TABLE_OPT_KEY`、`HIVE_PARTITION_FIELDS_OPT_KEY`、`HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY` を渡すのは必須ではなくなります。これらの値は、Hudi テーブル名とパーティションフィールドから推測できます。
  - `KEYGENERATOR_CLASS_OPT_KEY` を渡すのは必須ではなく、より単純な `SimpleKeyGenerator` および `ComplexKeyGenerator` のケースから推測できます。

## Hudi Caveats

- Hudi では、読み取り時マージ (MoR) テーブルおよびブートストラップテーブルの Hive でのベクトル化された実行をサポートしていません。例えば、`hive.vectorized.execution.enabled` が `true` に設定されている場合、Hudi リアルタイムテーブルで `count(*)` は失敗します。回避策として、`hive.vectorized.execution.enabled` を `false` に設定して、ベクトル化された読み取りを無効にすることができます。
- マルチライターサポートは、Hudi ブートストラップ機能とは互換性がありません。
- Flink Streamer と Flink SQL は、このリリースでは実験的な機能です。これらの機能は、実稼働環境での使用はお勧めしません。

## 変更点、機能強化、解決した問題

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

- 以前は、マルチマスタークラスターでリソースマネージャーを手動で再起動すると、Zookeeper などの Amazon EMR のクラスター上のデーモンが、Zookeeper `znode` ファイル内の以前に廃止された、または失われたすべてのノードを再ロードしていました。これにより、特定の状況でデフォルトの制限を超えることがありました。Amazon EMR では、1 時間以上経過した、廃止された、または失われたノードレコードが Zookeeper ファイルから削除されるようになり、内部制限も引き上げられました。

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Apache YARN タイムラインサーバーのバージョン 1 および 1.5 のパフォーマンスの問題を修正するためのクラスターの設定

Apache YARN タイムラインサーバーのバージョン 1 および 1.5 では、特に Amazon EMR のデフォルト設定である `yarn.resourcemanager.system-metrics-publisher.enabled=true` を使用する場合に、非常にアクティブで大規模な EMR クラスターでパフォーマンスの問題が発生する可能性があります。オープンソースの YARN タイムラインサーバー v2 では、YARN タイムラインサーバーのスケラビリティに関連するパフォーマンスの問題が解決されています。

この問題の他の回避策には、次のものがあります。

- `yarn-site.xml` での `yarn.resourcemanager.system-metrics-publisher.enabled=false` の設定。
- クラスターの作成時にこの問題の修正を有効にします (以下を参照)。

次の Amazon EMR リリースには、この YARN タイムラインサーバーのパフォーマンス問題の修正が含まれています。

EMR 5.30.2、5.31.1、5.32.1、5.33.1、5.34.x、6.0.1、6.1.1、6.2.1、6.3.1、6.4.x

上記に示されている Amazon EMR リリースで修正を有効にするには、[aws emr create-cluster コマンドのパラメータ](#) `--configurations file://./configurations.json` で渡される設定の JSON ファイルで、以下のプロパティを `true` に設定します。または、[再構成コンソール UI](#) を使用して修正を有効にします。

configurations.json ファイルの内容の例

```
[
{
  "Classification": "yarn-site",
  "Properties": {
    "yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
      "true",
    "yarn.resourcemanager.system-metrics-publisher.enabled": "true"
  },
  "Configurations": []
}
]
```

- WebHDFS と HttpFS サーバーはデフォルトで無効になっています。Hadoop 設定 `dfs.webhdfs.enabled` を使用して WebHDFS を再度有効にすることができます。HttpFS サーバーを起動するには、`sudo systemctl start hadoop-httpfs` を使用します。
- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPC ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (`$region` を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。[Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)
- Hive: 最後のジョブに HDFS でスクラッチディレクトリを使用できるようにすると、書き込みクエリのパフォーマンスが向上します。最後のジョブの一時データは Amazon S3 ではなく HDFS に書き込まれ、データは Amazon S3 デバイス間ではなく HDFS から最後のテーブルの場所 (Amazon S3) に移動されるため、パフォーマンスが向上します。
- Hive: Glue メタストアパーティショニングにより、クエリのコンパイル時間が最大 2.5 倍に改善されています。
- デフォルトでは、組み込みの UDF が Hive から Hive メタストアサーバーに渡されると、Glue は限定された式演算子しかサポートしていないため、組み込みの UDF のサブセットのみが Glue メタストアに渡されます。`hive.glue.partition.pruning.client=true`

を設定した場合、パーティションのプルーニングはすべてクライアント側で行われます。hive.glue.partition.pruning.server=true を設定した場合、パーティションのプルーニングはすべてサーバー側で行われます。

## 既知の問題

- Apache Hadoop HttpFS サーバーはデフォルトで無効になるため、Hue クエリは Amazon EMR 6.4.0 では機能しません。Amazon EMR 6.4.0 で Hue を使用するには、`sudo systemctl start hadoop-httpfs` を使用して Amazon EMR プライマリノードで HttpFS サーバーを手動で起動するか、または [Amazon EMR のステップを使用](#) します。
- Livy ユーザー偽装で使用される Amazon EMR Notebooks 機能は、HttpFS がデフォルトで無効になっているため機能しません。この場合、EMR notebooks は Livy 偽装が有効になっているクラスターに接続できません。回避策は、`sudo systemctl start hadoop-httpfs` を使用して EMR notebooks をクラスターに接続する前に、HttpFS サーバーを起動することです。
- Amazon EMR バージョン 6.4.0 では、Phoenix は Phoenix コネクタコンポーネントをサポートしていません。
- Apache Oozie で Spark アクションを使用するには、以下の設定を Oozie workflow.xml ファイルに追加する必要があります。そうしないと、Hadoop や EMRFS などのいくつかの重要なライブラリが、Oozie が起動する Spark エグゼキューターのクラスパスから失われてしまいます。

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-.`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## 6.4.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。

コンポーネント	Version	説明
emr-notebook-env	1.3.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.18.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.1.0	EMR S3Select コネクタ
emrfs	2.47.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.13.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.13.1	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	3.2.1-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httfs-server	3.2.1-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	3.2.1-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.4.4-amzn-0	リージョンの調整および管理コマンドの実行を行う HBase クラスタ用のサービス。
hbase-region-server	2.4.4-amzn-0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.4.4-amzn-0	HBase コマンドラインクライアント。
hbase-rest-server	2.4.4-amzn-0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.4.4-amzn-0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	3.1.2-amzn-5	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-5	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.2-amzn-5	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-5	Hive コマンドラインクライアント。



コンポーネント	Version	説明
hive-hbase	3.1.2-amzn-5	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-5	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-5	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.8.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.8.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-trino	0.8.0-amzn-0	Trino を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.8.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.9.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.8.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.254.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.254.1-amzn-0	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
presto-client	0.254.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
trino-coordinator	359	trino-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
trino-worker	359	いくつかのクエリを実行するサービス。
trino-client	359	Trino サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Trino コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.1.2-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.1.2-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	3.1.2-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.1.2-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	0.4.1	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.5.7	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.5.7	ZooKeeper コマンドラインクライアント。

## 6.4.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

### emr-6.4.0 分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

分類	説明	再設定アクション
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.



分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.

分類	説明	再設定アクション
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.

分類	説明	再設定アクション
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
trino-log	Trino の log.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-config	Trino の config.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)

分類	説明	再設定アクション
trino-password-authenticator	Trino の password-authenticator.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-env	Trino の trino-env.sh ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-node	Trino の node.properties ファイルの値を変更します。	Not available.
trino-connector-blackhole	Trino の blackhole.properties ファイルの値を変更します。	Not available.
trino-connector-cassandra	Trino の cassandra.properties ファイルの値を変更します。	Not available.
trino-connector-hive	Trino の hive.properties ファイルの値を変更します。	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino の jmx.properties ファイルの値を変更します。	Not available.
trino-connector-kafka	Trino の kafka.properties ファイルの値を変更します。	Not available.
trino-connector-localfile	Trino の localfile.properties ファイルの値を変更します。	Not available.
trino-connector-memory	Trino の memory.properties ファイルの値を変更します。	Not available.
trino-connector-mongodb	Trino の mongodb.properties ファイルの値を変更します。	Not available.
trino-connector-mysql	Trino の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
trino-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
trino-connector-raptor	Trino の raptor.properties ファイルの値を変更します。	Not available.
trino-connector-redis	Trino の redis.properties ファイルの値を変更します。	Not available.
trino-connector-redshift	Trino の redshift.properties ファイルの値を変更します。	Not available.
trino-connector-tpch	Trino の tpch.properties ファイルの値を変更します。	Not available.
trino-connector-tpcds	Trino の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.

分類	説明	再設定アクション
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 6.3.1

### 6.3.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
AWS SDK for Java	1.11.977	1.11.977	1.11.880	1.11.880
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.245.1	0.245.1	0.238.3	0.238.3
Spark	3.1.1	3.1.1	3.0.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	350	350	343	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

### 6.3.1 リリースノート

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

#### 変更、拡張、解決した問題

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのへ

ルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。

- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケーリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケーリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPCE ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (`$region` を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。 [Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)

## 既知の問題

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに

注意してください: !"#%&'()\*+,-. 詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の spark.sql.sources.fastS3PartitionDiscovery.enabled 設定を false にセットします。

### 6.3.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。

コンポーネント	Version	説明
emr-notebook-env	1.2.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.18.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.1.0	EMR S3Select コネクタ
emrfs	2.46.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.12.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.12.1	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	3.2.1-amzn-3.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-3.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-3.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-3.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-3.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-3.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-3.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-3.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-3.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-3.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。



コンポーネント	Version	説明
hadoop-yarn-timeline-server	3.2.1-amzn-3.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.2.6-amzn-1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.2.6-amzn-1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.2.6-amzn-1	HBase コマンドラインクライアント。
hbase-rest-server	2.2.6-amzn-1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.2.6-amzn-1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	3.1.2-amzn-4	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-4	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.2-amzn-4	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-4	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-hbase	3.1.2-amzn-4	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-4	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-4	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.7.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.7.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-prestosql	0.7.0-amzn-0	PrestoSQL を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.7.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.9.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.2.2	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.7.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.0.0-HBase-2.0	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.0.0-HBase-2.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.245.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.245.1-amzn-0	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
presto-client	0.245.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
prestoql-coordinator	350	prestoql-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
prestoql-worker	350	いくつかのクエリを実行するサービス。
prestoql-client	350	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.1.1-amzn-0.1	Spark コマンドラインクライアント。
spark-history-server	3.1.1-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	3.1.1-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.1.1-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	0.4.1	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 6.3.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-6.3.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

分類	説明	再設定アクション
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.



分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.

分類	説明	再設定アクション
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.

分類	説明	再設定アクション
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
prestoql-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestoql-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)

分類	説明	再設定アクション
prestosql-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-node	PrestoSQL の node.properties ファイルの値を変更します。	Not available.
prestosql-connector-blackhole	PrestoSQL の blackhole.properties ファイルの値を変更します。	Not available.
prestosql-connector-cassandra	PrestoSQL の cassandra.properties ファイルの値を変更します。	Not available.
prestosql-connector-hive	PrestoSQL の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	PrestoSQL の jmx.properties ファイルの値を変更します。	Not available.
prestosql-connector-kafka	PrestoSQL の kafka.properties ファイルの値を変更します。	Not available.
prestosql-connector-localfile	PrestoSQL の localfile.properties ファイルの値を変更します。	Not available.
prestosql-connector-memory	PrestoSQL の memory.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
prestosql-connector-mongodb	PrestoSQL の mongodb.properties ファイルの値を変更します。	Not available.
prestosql-connector-mysql	PrestoSQL の mysql.properties ファイルの値を変更します。	Not available.
prestosql-connector-postgresql	PrestoSQL の postgresql.properties ファイルの値を変更します。	Not available.
prestosql-connector-raptor	PrestoSQL の raptor.properties ファイルの値を変更します。	Not available.
prestosql-connector-redis	PrestoSQL の redis.properties ファイルの値を変更します。	Not available.
prestosql-connector-redshift	PrestoSQL の redshift.properties ファイルの値を変更します。	Not available.
prestosql-connector-tpch	PrestoSQL の tpch.properties ファイルの値を変更します。	Not available.
prestosql-connector-tpcds	PrestoSQL の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j .properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 6.3.0

### 6.3.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
AWS SDK for Java	1.11.977	1.11.880	1.11.880	1.11.828
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.245.1	0.238.3	0.238.3	0.232
Spark	3.1.1	3.0.1	3.0.1	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
Trino (PrestoSQL)	350	343	343	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 6.3.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.3.0 に関する情報が含まれています。6.2.0 からの変更が含まれています。

初回リリース日: 2021 年 5 月 12 日

最終更新日: 2021 年 8 月 9 日

### サポートされているアプリケーション

- AWS SDK for Java バージョン 1.11.977
- CloudWatch シンクバージョン 2.1.0
- DynamoDB Connector バージョン 4.16.0
- EMRFS バージョン 2.46.0
- Amazon EMR Goodies バージョン 3.2.0
- Amazon EMR Kinesis Connector バージョン 3.5.0
- Amazon EMR Record Server バージョン 2.0.0
- Amazon EMR Scripts バージョン 2.5.0
- Flink バージョン 1.12.1
- Ganglia バージョン 3.7.2
- AWS Glue Hive メタストアクライアントバージョン 3.2.0
- Hadoop バージョン 3.2.1-amzn-3
- HBase バージョン 2.2.6-amzn-1
- HBase-operator-tools 1.0.0
- HCatalog バージョン 3.1.2-amzn-0

- Hive バージョン 3.1.2-amzn-4
- Hudi バージョン 0.7.0-amzn-0
- Hue バージョン 4.9.0
- Java JDK バージョン Corretto-8.282.08.1 (ビルド 1.8.0\_282-b08)
- JupyterHub バージョン 1.2.0
- Livy バージョン 0.7.0-incubating
- MXNet バージョン 1.7.0
- Oozie バージョン 5.2.1
- Phoenix バージョン 5.0.0
- Pig バージョン 0.17.0
- Presto バージョン 0.245.1-amzn-0
- PrestoSQL バージョン 350
- Apache Ranger KMS (マルチマスター透過的暗号化) バージョン 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.1.0
- SageMaker Spark SDK バージョン 1.4.1
- Scala バージョン 2.12.10 (OpenJDK 64 ビットサーバー VM、Java 1.8.0\_282)
- Spark バージョン 3.1.1-amzn-0
- spark-rapids 0.4.1
- Sqoop バージョン 1.4.7
- TensorFlow バージョン 2.4.1
- tez バージョン 0.9.2
- Zeppelin バージョン 0.9.0
- Zookeeper バージョン 3.4.14
- コネクタおよびドライバー: DynamoDB Connector 4.16.0

## 新機能

- Amazon EMR は Amazon S3 Access Points をサポートしています。これは、共有データレイクへのアクセスを簡単に管理できる Amazon S3 の機能です。Amazon S3 アクセスポイントのエイリアスを使用すると、Amazon EMR で大規模なデータアクセスを簡素化できます。Amazon EMR が利用可能なすべてのリージョンで、Amazon EMR のすべてのバージョンで Amazon S3 アクセス

ポイントを追加料金なしで使用できます。AWS Amazon S3 Access Points および Access Point のエイリアスの詳細については、「Amazon S3 ユーザーガイド」の「[アクセスポイントでのバケット形式のエイリアスの使用](#)」を参照してください。

- 新しい DescribeReleaseLabel API パラメータおよび ListReleaseLabel API パラメータは、Amazon EMR リリースラベルの詳細を提供します。API リクエストが実行されているリージョンで利用可能なリリースをプログラムで一覧表示でき、特定の Amazon EMR リリースラベルで使用可能なアプリケーションをプログラムで一覧表示できます。リリースラベルのパラメータにより、Spark などの指定したアプリケーションをサポートしている Amazon EMR リリースも一覧表示されます。この情報は、Amazon EMR クラスターをプログラムで起動するために使用できます。たとえば、ListReleaseLabel の結果から最新のリリースバージョンを使用してクラスターを起動できます。詳細については、「Amazon EMR API リファレンス [ListReleaseLabels](#)」の [DescribeReleaseLabel](#) 「」および 「」を参照してください。
- Amazon EMR 6.3.0 では、Apache Ranger とネイティブに統合されるクラスターを起動できます。Apache Ranger は、Hadoop プラットフォーム全体で包括的なデータセキュリティを有効化、監視、管理するためのオープンソースフレームワークです。詳細については、「[Apache Ranger](#)」を参照してください。ネイティブ統合により、独自の Apache Ranger を導入して、Amazon EMR できめ細かなデータアクセス制御を実施できます。「Amazon EMR 管理ガイド」の「[Amazon EMR と Apache Ranger を統合する](#)」を参照してください。
- スコープ管理ポリシー: AWS ベストプラクティスに合わせて、Amazon EMR では v2 EMR スコープのデフォルト管理ポリシーを廃止されるポリシーの置き換えとして導入しました。「[Amazon EMR 管理ポリシー](#)」を参照してください。
- インスタンスメタデータサービス (IMDS) V2 サポートステータス: Amazon EMR 6.2 以降では、Amazon EMR コンポーネントはすべての IMDS 呼び出しに IMDSv2 を使用します。アプリケーションコードでの IMDS 呼び出しでは、IMDSv1 と IMDSv2 の両方を使用するか、または、セキュリティを強化するために IMDSv2 のみを使用するように IMDS を設定できます。以前の Amazon EMR 6.x リリースで IMDSv1 を無効にすると、クラスターの起動が失敗します。

## 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのへ



ルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。

- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケーリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケーリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Spark SQL UI の explain モードのデフォルトは [Spark 3.1](#) で extended から formatted に変更されました。Amazon EMR では、それを extended に戻し、Spark SQL UI に論理プラン情報を含めるようになっています。spark.sql.ui.explainMode を formatted に設定すると、これを元に戻すことができます。
- 以下のコミットは Spark マスターブランチからバックポートされました。
  - [\[SPARK-34752\]](#)[ビルド] CVE-2020-27223 に対処するために Jetty を 9.4.37 にバンプ。
  - [\[SPARK-34534\]](#) FetchShuffleBlocks を使用してブロックを取得するときの blockIds の順序を修正しました。
  - [\[SPARK-34681\]](#) [SQL] 非等化条件で左側を構築するときの、完全外部のシャッフルされたハッシュ結合のバグを修正。
  - [\[SPARK-34497\]](#) [SQL] JVM セキュリティコンテキストの変更を復元するための、組み込みの JDBC 接続プロバイダーを修正。

- Nvidia Spark RAPIDS プラグインとの相互運用性を向上させるために、アダプティブクエリ実行を無効にして Nvidia Spark RAPIDS を使用するとき動的パーティションプルーフングがトリガーされない問題に対処する回避策を追加しました。「[RAPIDS の問題 #1378](#)」および「[RAPIDS の問題 #1386](#)」を参照してください。新しい設定 `spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse` の詳細については、「[RAPIDS の問題 #1386](#)」を参照してください。
- ファイル出力コミッターのデフォルトアルゴリズムは、オープンソースの Spark 3.1 の v2 アルゴリズムから v1 アルゴリズムに変更されました。詳細については、「[Amazon EMR による Spark パフォーマンスの最適化 - ダイナミックパーティションプルーフング](#)」を参照してください。
- パフォーマンスの低下を防ぐため、Amazon EMR は以前の Amazon EMR 6.x リリースで使用されていたデフォルトである v2 アルゴリズムに戻りました。オープンソースの Spark 3.1 の動作を復元するには、`spark.hadoop.mapreduce.fileoutputcommitter.algorithm.version` を 1 に設定します。オープンソースの Spark では、ファイル出力コミッターのアルゴリズム v2 のタスクコミットがアトミックではなく、場合によっては出力データの正確性の問題を引き起こす可能性があるため、この変更が行われました。ただし、アルゴリズム v1 のタスクコミットもアトミックではありません。場合によっては、タスクコミットに名前変更の前に実行された削除が含まれます。これにより、サイレントデータの正確性の問題が発生する可能性があります。
- 以前の Amazon EMR リリースでマネージドスケールリングの問題を修正し、アプリケーションの障害率が大幅に低下するように改善しました。
- 新しいクラスターごとに AWS Java SDK バンドルをインストールしました。これは、個々のコンポーネント jar ではなく、すべてのサービス SDK とその依存関係を含む単一の jar です。詳細については、「[Java SDK Bundled Dependency](#)」を参照してください。

## 既知の問題

- Amazon EMR 6.3.0 および 6.2.0 のプライベートサブネットクラスターで、Ganglia ウェブ UI にアクセスできません。「アクセス拒否 (403)」というエラーが表示されます。Spark、Hue、Zeppelin、Livy JupyterHub、Tez などの他のウェブ UIs は正常に動作しています。パブリックサブネットクラスターでの Ganglia ウェブ UI アクセスも正常に動作します。この問題を解決するには、`sudo systemctl restart httpd` を使用してプライマリノードで httpd サービスを再起動します。この問題は、Amazon EMR 6.4.0 で修正されています。
- AWS Glue Data Catalog が有効になっている場合、Spark を使用して null 文字列の場所 URI を持つ AWS Glue DB にアクセスすると、失敗する可能性があります。これは以前の Amazon EMR リリースで発生しますが、SPARK-31709 (<https://issues.apache.org/jira/browse/SPARK-31709>) ではより多くの場合に当てはまります。例えば、場所 URI が null 文字列であるデフォルトの

AWS Glue DB 内にテーブルを作成すると、は「空の文字列からパスを作成できません」というメッセージで `spark.sql("CREATE TABLE mytest (key string) location '/table_path';")` 失敗します。これを回避するには、AWS Glue データベースの場所 URI を手動で設定し、Spark を使用してこれらのデータベース内にテーブルを作成します。

- Amazon EMR 6.3.0 では、PrestoSQL がバージョン 343 からバージョン 350 にアップグレードされました。このバージョン変更に関連して、オープンソースからのセキュリティ関連の変更が 2 つあります。ファイルベースのカatalog アクセス制御は、テーブル、スキーマ、またはセッションのプロパティのルールが定義されていない場合、deny から allow に変更されています。また、ファイルベースのシステムアクセス制御は、Catalog ルールが定義されていないファイルをサポートするように変更されています。この場合、Catalog へのすべてのアクセスが許可されます。

詳細については、「[Release 344 \(9 Oct 2020\)](#)」を参照してください。

- Hadoop ユーザーディレクトリ (/home/hadoop) はすべての人が読み取り可能であることに注意してください。Hive などのフレームワークによる読み取りアクセスを許可する Unix 755 (drwxr-xr-x) ディレクトリのアクセス許可があります。/home/hadoop とそのサブディレクトリにファイルを配置できますが、機密情報を保護するために、これらのディレクトリに対するアクセス許可に注意してください。
- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

## コマンドラインから ulimit を明示的に設定する

1. `/etc/systemd/system/instance-controller.service` を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

## ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

**⚠ Important**

Amazon Linux または Amazon Linux 2 Amazon マシンイメージ (AMI) を実行する EMR クラスターは、デフォルトの Amazon Linux 動作を使用します。再起動が必要な重要な脆弱なカーネル更新が自動的にダウンロードされてインストールされることはありません。これは、デフォルトの Amazon Linux AMI を実行している他の Amazon EC2 インスタンスと同じ動作です。Amazon EMR リリースが利用可能になった後に、再起動が必要な新しい Amazon Linux ソフトウェアアップデート (カーネル、NVIDIA、CUDA のアップデートなど) が使用可能になった場合、デフォルトの AMI を実行する EMR クラスターインスタンスで、それらの更新が自動的にダウンロードされてインストールされることはありません。カーネルの更新を取得するには、[Amazon EMR AMI をカスタマイズ](#)して、[最新の Amazon Linux AMI を使用](#)できるようにします。

- Apache Oozie で Spark アクションを使用するには、以下の設定を Oozie workflow.xml ファイルに追加する必要があります。そうしないと、Hadoop や EMRFS などのいくつかの重要なライブラリが、Oozie が起動する Spark エグゼキューターのクラスパスから失われてしまいます。

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、s3://bucket/table/p=a は s3://bucket/table/p=a b のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の spark.sql.sources.fastS3PartitionDiscovery.enabled 設定を false にセットします。

### 6.3.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-notebook-env</code>	1.2.0	Jupyter エンタープライズゲートウェイを含む <code>emr</code> ノートブック用 Conda env
<code>emr-s3-dist-cp</code>	2.18.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	2.1.0	EMR S3Select コネクタ

コンポーネント	Version	説明
emrfs	2.46.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.12.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.12.1	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-3	HDFS コマンドラインクライアントとライブラリ



コンポーネント	Version	説明
hadoop-hdfs-namenode	3.2.1-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-3	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.2.6-amzn-1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。



コンポーネント	Version	説明
hbase-region-server	2.2.6-amzn-1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.2.6-amzn-1	HBase コマンドラインクライアント。
hbase-rest-server	2.2.6-amzn-1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.2.6-amzn-1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	3.1.2-amzn-4	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-4	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.2-amzn-4	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-4	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-4	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-4	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。

コンポーネント	Version	説明
hive-server2	3.1.2-amzn-4	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.7.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.7.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-prestosql	0.7.0-amzn-0	PrestoSQL を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.7.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.9.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.2.2	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.7.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。

コンポーネント	Version	説明
mariadb-server	5.5.68+	MariaDB データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.0.0-HBase-2.0	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.0.0-HBase-2.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.245.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.245.1-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.245.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。

コンポーネント	Version	説明
prestoql-coordinator	350	prestoql-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
prestoql-worker	350	いくつかのクエリを実行するサービス。
prestoql-client	350	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.1.1-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.1.1-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.1.1-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.1.1-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。

コンポーネント	Version	説明
spark-rapids	0.4.1	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 6.3.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-6.3.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.



分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.

分類	説明	再設定アクション
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.

分類	説明	再設定アクション
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
prestoql-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestoql-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)

分類	説明	再設定アクション
prestosql-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-node	PrestoSQL の node.properties ファイルの値を変更します。	Not available.
prestosql-connector-blackhole	PrestoSQL の blackhole.properties ファイルの値を変更します。	Not available.
prestosql-connector-cassandra	PrestoSQL の cassandra.properties ファイルの値を変更します。	Not available.
prestosql-connector-hive	PrestoSQL の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	PrestoSQL の jmx.properties ファイルの値を変更します。	Not available.
prestosql-connector-kafka	PrestoSQL の kafka.properties ファイルの値を変更します。	Not available.
prestosql-connector-localfile	PrestoSQL の localfile.properties ファイルの値を変更します。	Not available.
prestosql-connector-memory	PrestoSQL の memory.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
prestosql-connector-mongodb	PrestoSQL の mongodb.properties ファイルの値を変更します。	Not available.
prestosql-connector-mysql	PrestoSQL の mysql.properties ファイルの値を変更します。	Not available.
prestosql-connector-postgresql	PrestoSQL の postgresql.properties ファイルの値を変更します。	Not available.
prestosql-connector-raptor	PrestoSQL の raptor.properties ファイルの値を変更します。	Not available.
prestosql-connector-redis	PrestoSQL の redis.properties ファイルの値を変更します。	Not available.
prestosql-connector-redshift	PrestoSQL の redshift.properties ファイルの値を変更します。	Not available.
prestosql-connector-tpch	PrestoSQL の tpch.properties ファイルの値を変更します。	Not available.
prestosql-connector-tpcds	PrestoSQL の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j .properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie and HiveServer2.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 6.2.1

### 6.2.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
AWS SDK for Java	1.11.880	1.11.880	1.11.828	1.11.828
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5	2.2.5

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.238.3	0.238.3	0.232	0.232
Spark	3.0.1	3.0.1	3.0.0	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
Trino (PrestoSQL)	343	343	338	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 6.2.1 リリースノート

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

### 変更、拡張、解決した問題

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース

5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。

- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPC ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (`$region` を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。[Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)

## 既知の問題

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## 6.2.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.0.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.16.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.0.0	EMR S3Select コネクタ
emrfs	2.44.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。



コンポーネント	Version	説明
flink-client	1.11.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.11.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-2.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-2.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-2.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-2.1	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-hdfs-journalnode	3.2.1-amzn-2.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-2.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-2.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-2.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-2.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-2.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-2.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.2.6-amzn-0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.2.6-amzn-0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.2.6-amzn-0	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	2.2.6-amzn-0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.2.6-amzn-0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	3.1.2-amzn-3	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-3	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.2-amzn-3	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-3	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-3	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-3	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-3	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hudi	0.6.0-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.6.0-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-prestosql	0.6.0-amzn-1	PrestoSQL を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.6.0-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.8.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.7.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64+	MariaDB データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.0.0-HBase-2.0	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.0.0-HBase-2.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.238.3-amzn-1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.238.3-amzn-1	いくつかのクエリを実行するサービス。
presto-client	0.238.3-amzn-1	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
prestosql-coordinator	343	prestosql-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
prestosql-worker	343	いくつかのクエリを実行するサービス。
prestosql-client	343	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.0.1-amzn-0.1	Spark コマンドラインクライアント。
spark-history-server	3.0.1-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.0.1-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.0.1-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	0.2.0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。

コンポーネント	Version	説明
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.3.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0-preview1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 6.2.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-6.2.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the



分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Not available.
flink-log4j	Flink log4j.properties の設定を変更します。	Not available.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Not available.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.

分類	説明	再設定アクション
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.

分類	説明	再設定アクション
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
prestoql-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestoql-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)

分類	説明	再設定アクション
prestosql-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-node	PrestoSQL の node.properties ファイルの値を変更します。	Not available.
prestosql-connector-blackhole	PrestoSQL の blackhole.properties ファイルの値を変更します。	Not available.
prestosql-connector-cassandra	PrestoSQL の cassandra.properties ファイルの値を変更します。	Not available.
prestosql-connector-hive	PrestoSQL の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	PrestoSQL の jmx.properties ファイルの値を変更します。	Not available.
prestosql-connector-kafka	PrestoSQL の kafka.properties ファイルの値を変更します。	Not available.
prestosql-connector-localfile	PrestoSQL の localfile.properties ファイルの値を変更します。	Not available.
prestosql-connector-memory	PrestoSQL の memory.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
prestosql-connector-mongodb	PrestoSQL の mongodb.properties ファイルの値を変更します。	Not available.
prestosql-connector-mysql	PrestoSQL の mysql.properties ファイルの値を変更します。	Not available.
prestosql-connector-postgresql	PrestoSQL の postgresql.properties ファイルの値を変更します。	Not available.
prestosql-connector-raptor	PrestoSQL の raptor.properties ファイルの値を変更します。	Not available.
prestosql-connector-redis	PrestoSQL の redis.properties ファイルの値を変更します。	Not available.
prestosql-connector-redshift	PrestoSQL の redshift.properties ファイルの値を変更します。	Not available.
prestosql-connector-tpch	PrestoSQL の tpch.properties ファイルの値を変更します。	Not available.
prestosql-connector-tpcds	PrestoSQL の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j .properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 6.2.0

### 6.2.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
AWS SDK for Java	1.11.880	1.11.828	1.11.828	1.11.711
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.5	2.2.5	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
Hudi	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1
Hue	4.8.0	4.7.1	4.7.1	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	-
Presto	0.238.3	0.232	0.232	0.230
Spark	3.0.1	3.0.0	3.0.0	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	-
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	338	338	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 6.2.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.2.0 に関する情報が含まれています。6.1.0 からの変更が含まれています。

初回リリース日: 2020 年 12 月 9 日

最終更新日: 2021 年 10 月 4 日

### サポートされているアプリケーション

- AWS SDK for Java バージョン 1.11.828
- emr-record-server バージョン 1.7.0
- Flink バージョン 1.11.2
- Ganglia バージョン 3.7.2
- Hadoop バージョン 3.2.1-amzn-1
- HBase バージョン 2.2.6-amzn-0
- HBase operator-tools 1.0.0
- HCatalog バージョン 3.1.2-amzn-0
- Hive バージョン 3.1.2-amzn-3
- Hudi バージョン 0.6.0-amzn-1
- Hue バージョン 4.8.0
- JupyterHub バージョン 1.1.0
- Livy バージョン 0.7.0
- MXNet バージョン 1.7.0
- Oozie バージョン 5.2.0
- Phoenix バージョン 5.0.0
- Pig バージョン 0.17.0
- Presto バージョン 0.238.3-amzn-1
- PrestoSQL バージョン 343
- Spark バージョン 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow バージョン 2.3.1
- Zeppelin バージョン 0.9.0-preview1



- Zookeeper バージョン 3.4.14
- コネクタおよびドライバー: DynamoDB Connector 4.16.0

## 新機能

- HBase: コミットフェーズでの名前変更を削除し、永続的 HFile トラッキングを追加しました。「Amazon EMR リリース ガイド」の「[永続的 HFile トラッキング](#)」を参照してください。
- HBase: 「[Create a config that forces to cache blocks on compaction](#)」をバックポートしました。
- PrestoDB: ダイナミックパーティションプルーニングの改善。ルールベースの結合順序は、パーティション分割されていないデータに対して機能します。
- スコープ管理ポリシー: AWS ベストプラクティスに合わせて、Amazon EMR では v2 EMR スコープのデフォルト管理ポリシーを廃止されるポリシーの置き換えとして導入しました。「[Amazon EMR 管理ポリシー](#)」を参照してください。
- インスタンスメタデータサービス (IMDS) V2 サポートステータス: Amazon EMR 6.2 以降では、Amazon EMR コンポーネントはすべての IMDS 呼び出しに IMDSv2 を使用します。アプリケーションコードでの IMDS 呼び出しでは、IMDSv1 と IMDSv2 の両方を使用するか、または、セキュリティを強化するために IMDSv2 のみを使用するように IMDS を設定できます。以前の Amazon EMR 6.x リリースで IMDSv1 を無効にすると、クラスターの起動が失敗します。

## 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールアップリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。

- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Spark: Spark ランタイムのパフォーマンスが向上しました。

## 既知の問題

- Amazon EMR 6.2 の EMR 6.2.0 の `/etc/cron.d/libinstance-controller-java` file に正しくないアクセス許可が設定されている。ファイルに対するアクセス許可は、644 (`-rw-r--r--`) であるべきときに 645 (`-rw-r--r-x`) になります。その結果、Amazon EMR バージョン 6.2 ではインスタンス状態ログが記録されず、`/emr/instance-logs` ディレクトリは空になります。この問題は、Amazon EMR 6.3.0 以降で修正されています。

この問題を回避するには、クラスター起動時に以下のスクリプトをブートストラップアクションとして実行します。

```
#!/bin/bash
sudo chmod 644 /etc/cron.d/libinstance-controller-java
```

- Amazon EMR 6.2.0 および 6.3.0 のプライベートサブネットクラスターで、Ganglia ウェブ UI にアクセスできません。「アクセス拒否 (403)」というエラーが表示されます。Spark、Hue、Zeppelin、Livy JupyterHub、Tez などの他のウェブ UIs は正常に動作しています。パブリックサブネットクラスターでの Ganglia ウェブ UI アクセスも正常に動作します。この問題を解決するには、`sudo systemctl restart httpd` を使用してプライマリノードで httpd サービスを再起動します。この問題は、Amazon EMR 6.4.0 で修正されています。
- Amazon EMR 6.2.0 で、httpd が連続して失敗し、Ganglia が利用できないという問題があります。「サーバーに接続できません」というエラーが表示されます。この問題が発生している、既に実行されているクラスターを修正するには、クラスターのプライマリノードに SSH 接続して、/

etc/httpd/conf/httpd.conf にあるファイル httpd.conf に行 Listen 80 を追加します。この問題は、Amazon EMR 6.3.0 で修正されています。

- セキュリティ設定を使用している場合、EMR 6.2.0 クラスターで HTTPD が失敗します。これにより、Ganglia ウェブアプリケーションのユーザーインターフェイスが使用できなくなります。Ganglia ウェブアプリケーションのユーザーインターフェイスにアクセスするには、クラスターのプライマリノード上の /etc/httpd/conf/httpd.conf ファイルに Listen 80 を追加します。クラスターへの接続の詳細については、「[SSH を使用してプライマリノードに接続する](#)」を参照してください。

EMR Notebooks では、セキュリティ構成を使用する場合、EMR 6.2.0 クラスターとの接続の確立にも失敗します。ノートブックは、カーネルのリストと Spark ジョブの送信に失敗します。代わりに、別のバージョンの Amazon EMR で EMR Notebooks を使用することをお勧めします。

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

#### Important

Amazon EMR 6.1.0 および 6.2.0 には、Hudi の挿入、アップサート、および削除操作すべてに大きな影響を及ぼす可能性があるパフォーマンスの問題が含まれています。Amazon EMR 6.1.0 または 6.2.0 で Hudi を使用する予定がある場合は、AWS サポートに連絡してパッチが適用された Hudi RPM を取得する必要があります。

#### Important

Amazon Linux または Amazon Linux 2 Amazon マシンイメージ (AMI) を実行する EMR クラスターは、デフォルトの Amazon Linux 動作を使用します。再起動が必要な重要かつクリティカルなカーネル更新が自動的にダウンロードされてインストールされることはあり

ません。これは、デフォルトの Amazon Linux AMI を実行している他の Amazon EC2 インスタンスと同じ動作です。Amazon EMR リリースが利用可能になった後に、再起動が必要な新しい Amazon Linux ソフトウェアアップデート (カーネル、NVIDIA、CUDA のアップデートなど) が使用可能になった場合、デフォルトの AMI を実行する EMR クラスターインスタンスで、それらの更新が自動的にダウンロードされてインストールされることはありません。カーネルの更新を取得するには、[Amazon EMR AMI をカスタマイズ](#)して、[最新の Amazon Linux AMI を使用](#)できるようにします。

- Amazon EMR 6.2.0 の Maven アーティファクトは公開されません。これらは Amazon EMR の将来のリリースで公開される予定です。
- HBase ストアファイルシステムテーブルを使用した永続的な HFile トラッキングは、HBase リージョンのレプリケーション機能をサポートしていません。HBase リージョンレプリケーションの詳細については、「[Timeline-consistent High Available Reads](#)」を参照してください。
- Amazon EMR 6.x と EMR 5.x の Hive バケット化バージョンの違い

EMR 5.x では OOS Apache Hive 2 を使用していますが、EMR 6.x では OOS Apache Hive 3 を使用しています。オープンソースの Hive2 ではバケット化バージョン 1 を使用していますが、オープンソースの Hive3 ではバケット化バージョン 2 を使用しています。Hive 2 (EMR 5.x) と Hive 3 (EMR 6.x) のこのバケット化バージョンの違いは、Hive のバケット化ハッシュ関数が異なることを意味します。以下の例を参照してください。

次の表は、それぞれ EMR 6.x および EMR 5.x で作成した例です。

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

EMR 6.x と EMR 5.x の両方に同じデータを挿入しています。

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

S3 の場所を確認すると、バケット化ファイル名が異なることが示されています。これは、ハッシュ関数が EMR 6.x (Hive 3) と EMR 5.x (Hive 2) で異なるためです。

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16          13 000025_0
2020-10-21 20:35:22          14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07          13 000066_0
2020-10-21 20:32:51          14 000072_0
```

EMR 6.x で、Hive CLI で次のコマンドを実行し、バージョンの違いを確認することもできます。バケット化バージョン 2 が返されることに注意してください。

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
  bucketing_version      2
...
```

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。



**Note**

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-.`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## 6.2.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオー

プンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.0.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.16.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.0.0	EMR S3Select コネクタ
emrfs	2.44.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.11.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.11.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。



コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-2	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-2	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	3.2.1-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.2.6-amzn-0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.2.6-amzn-0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.2.6-amzn-0	HBase コマンドラインクライアント。
hbase-rest-server	2.2.6-amzn-0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.2.6-amzn-0	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	3.1.2-amzn-3	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-3	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.2-amzn-3	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-3	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-3	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-3	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-3	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.6.0-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.6.0-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hudi-prestosql	0.6.0-amzn-1	PrestoSQL を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.6.0-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.8.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.7.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64+	MariaDB データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
opencv	4.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.0.0-HBase-2.0	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.0.0-HBase-2.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.238.3-amzn-1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.238.3-amzn-1	いくつかのクエリを実行するサービス。
presto-client	0.238.3-amzn-1	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
prestosql-coordinator	343	prestosql-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
prestosql-worker	343	いくつかのクエリを実行するサービス。
prestosql-client	343	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。

コンポーネント	Version	説明
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.0.1-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.0.1-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.0.1-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.0.1-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
spark-rapids	0.2.0	GPU で Apache Spark を高速化する Nvidia Spark RAPIDS プラグイン。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.3.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0-preview1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 6.2.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

### emr-6.2.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.



分類	説明	再設定アクション
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Not available.
flink-log4j	Flink log4j.properties の設定を変更します。	Not available.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Not available.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	This classification should not be reconfigured.
hdfs-env	HDFS 環境の値を変更します。	Restarts Hadoop HDFS ZKFC.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat server.
hive	Apache Hive 用の Amazon EMR-curated 設定。	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Not available.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.

分類	説明	再設定アクション
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.

分類	説明	再設定アクション
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
prestoql-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestoql-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)

分類	説明	再設定アクション
prestosql-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-node	PrestoSQL の node.properties ファイルの値を変更します。	Not available.
prestosql-connector-blackhole	PrestoSQL の blackhole.properties ファイルの値を変更します。	Not available.
prestosql-connector-cassandra	PrestoSQL の cassandra.properties ファイルの値を変更します。	Not available.
prestosql-connector-hive	PrestoSQL の hive.properties ファイルの値を変更します。	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	PrestoSQL の jmx.properties ファイルの値を変更します。	Not available.
prestosql-connector-kafka	PrestoSQL の kafka.properties ファイルの値を変更します。	Not available.
prestosql-connector-localfile	PrestoSQL の localfile.properties ファイルの値を変更します。	Not available.
prestosql-connector-memory	PrestoSQL の memory.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
prestosql-connector-mongodb	PrestoSQL の mongodb.properties ファイルの値を変更します。	Not available.
prestosql-connector-mysql	PrestoSQL の mysql.properties ファイルの値を変更します。	Not available.
prestosql-connector-postgresql	PrestoSQL の postgresql.properties ファイルの値を変更します。	Not available.
prestosql-connector-raptor	PrestoSQL の raptor.properties ファイルの値を変更します。	Not available.
prestosql-connector-redis	PrestoSQL の redis.properties ファイルの値を変更します。	Not available.
prestosql-connector-redshift	PrestoSQL の redshift.properties ファイルの値を変更します。	Not available.
prestosql-connector-tpch	PrestoSQL の tpch.properties ファイルの値を変更します。	Not available.
prestosql-connector-tpcds	PrestoSQL の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j .properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restart Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 6.1.1

### 6.1.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Oozie](#)、[Ph](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 6.1.1 リリースノート

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

### 変更、拡張、解決した問題

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールアップリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPC ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (\$region を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。 [Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)



### 6.1.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.3.0	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.14.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	3.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.14.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	2.0.0	EMR S3Select コネクタ
<code>emrfs</code>	2.42.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.11.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-1.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-1.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-1.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-1.1	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-hdfs-journalnode	3.2.1-amzn-1.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-1.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-1.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-1.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-1.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-1.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-1.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.2.5	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.2.5	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.2.5	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	2.2.5	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.2.5	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	3.1.2-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.2-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-2	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-2	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hudi	0.5.2-incubating-amzn-2	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.2-incubating-amzn-2	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-prestosql	0.5.2-incubating-amzn-2	PrestoSQL を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.5.2-incubating-amzn-2	Spark を Hudi で実行するためのバンドルライブラリ。
hue-server	4.7.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.6.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64+	MariaDB データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.3.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.0.0-HBase-2.0	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.0.0-HBase-2.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.232	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.232	いくつかのクエリを実行するサービス。
presto-client	0.232	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
prestosql-coordinator	338	prestosql-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
prestoql-worker	338	いくつかのクエリを実行するサービス。
prestoql-client	338	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.0.0-amzn-0.1	Spark コマンドラインクライアント。
spark-history-server	3.0.0-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.0.0-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.0.0-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	2.1.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0-preview1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 6.1.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-6.1.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。



分類	説明
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-env	HDFS 環境の値を変更します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive	Apache Hive 用の Amazon EMR-curated 設定。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。

分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
prestoql-log	Presto の log.properties ファイルの値を変更します。
prestoql-config	Presto の config.properties ファイルの値を変更します。

分類	説明
prestosql-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
prestosql-env	Presto の presto-env.sh ファイルの値を変更します。
prestosql-node	PrestoSQL の node.properties ファイルの値を変更します。
prestosql-connector-blackhole	PrestoSQL の blackhole.properties ファイルの値を変更します。
prestosql-connector-cassandra	PrestoSQL の cassandra.properties ファイルの値を変更します。
prestosql-connector-hive	PrestoSQL の hive.properties ファイルの値を変更します。
prestosql-connector-jmx	PrestoSQL の jmx.properties ファイルの値を変更します。
prestosql-connector-kafka	PrestoSQL の kafka.properties ファイルの値を変更します。
prestosql-connector-localfile	PrestoSQL の localfile.properties ファイルの値を変更します。
prestosql-connector-memory	PrestoSQL の memory.properties ファイルの値を変更します。
prestosql-connector-mongodb	PrestoSQL の mongodb.properties ファイルの値を変更します。
prestosql-connector-mysql	PrestoSQL の mysql.properties ファイルの値を変更します。
prestosql-connector-postgresql	PrestoSQL の postgresql.properties ファイルの値を変更します。

分類	説明
prestosql-connector-raptor	PrestoSQL の raptor.properties ファイルの値を変更します。
prestosql-connector-redis	PrestoSQL の redis.properties ファイルの値を変更します。
prestosql-connector-redshift	PrestoSQL の redshift.properties ファイルの値を変更します。
prestosql-connector-tpch	PrestoSQL の tpch.properties ファイルの値を変更します。
prestosql-connector-tpcds	PrestoSQL の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。



分類	説明
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 6.1.0

### 6.1.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Oozie](#)、[Phoenix](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 6.1.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.1.0 に関する情報が含まれています。6.0.0 からの変更が含まれています。

初回リリース日: 2020 年 9 月 4 日

最終更新日: 2020 年 10 月 15 日

サポートされているアプリケーション

- AWS SDK for Java バージョン 1.11.828
- Flink バージョン 1.11.0
- Ganglia バージョン 3.7.2
- Hadoop バージョン 3.2.1-amzn-1
- HBase バージョン 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog バージョン 3.1.2-amzn-0
- Hive バージョン 3.1.2-amzn-1
- Hudi バージョン 0.5.2-incubating
- Hue バージョン 4.7.1
- JupyterHub バージョン 1.1.0
- Livy バージョン 0.7.0
- MXNet バージョン 1.6.0
- Oozie バージョン 5.2.0
- Phoenix バージョン 5.0.0

- Presto バージョン 0.232
- PrestoSQL バージョン 338
- Spark バージョン 3.0.0-amzn-0
- TensorFlow バージョン 2.1.0
- Zeppelin バージョン 0.9.0-preview1
- Zookeeper バージョン 3.4.14
- コネクタおよびドライバー: DynamoDB Connector 4.14.0

## 新機能

- ARM インスタンスタイプは、Amazon EMR バージョン 5.30.0 以降および Amazon EMR バージョン 6.1.0 以降でサポートされています。
- M6g 汎用インスタンスタイプは、Amazon EMR バージョン 6.1.0 以降および Amazon EMR バージョン 5.30.0 以降でサポートされています。詳細については、「Amazon EMR 管理ガイド」の「[サポートされるインスタンスタイプ](#)」を参照してください。
- EC2 プレイACEMENTグループ機能は、複数のプライマリノードクラスターのオプションとして Amazon EMR バージョン 5.23.0 以降でサポートされています。現在、プレイACEMENTグループ機能ではプライマリノードタイプのみがサポートされており、SPREAD ストラテジーは、これらのプライマリノードに適用されます。SPREAD 戦略では、ハードウェア障害の発生時に複数のプライマリノードが失われるのを防ぐため、少数のインスタンスを別個の基盤となるハードウェア全体に配置します。詳細については、「Amazon EMR 管理ガイド」の「[EC2 プレイACEMENTグループと EMR の統合](#)」を参照してください。
- マネージドスケーリング - Amazon EMR バージョン 6.1.0 では、Amazon EMR Managed Scaling を有効にすることで、ワークロードに応じてクラスター内のインスタンスやユニットの数を自動的に増減できます。Amazon EMR は引き続きクラスターのメトリクスを評価し、クラスターのコストと速度を最適化するためのスケーリングを決定します。マネージドスケーリングは、Amazon EMR バージョン 5.30.0 以降 (6.0.0 を除く) でも使用できます。詳細については、「Amazon EMR 管理ガイド」の「[クラスターリソースのスケーリング](#)」を参照してください。
- PrestoSQL バージョン 338 は EMR 6.1.0 でサポートされています。詳細については、「[Presto](#)」を参照してください。
  - PrestoSQL は EMR 6.1.0 以降のバージョンでのみサポートされています。EMR 6.0.0 および EMR 5.x ではサポートされていません。

- アプリケーション名 Presto は、クラスターに PrestoDB をインストールするために引き続き使用されます。クラスターに PrestoDB をインストールするには、アプリケーション名 PrestoSQL を使用します。
- PrestoDB または PrestoSQL のいずれかをインストールできますが、両方を 1 つのクラスターにインストールすることはできません。クラスターの作成時に PrestoDB と PrestoSQL の両方を指定すると、検証エラーが発生し、クラスター作成リクエストは失敗します。
- PrestoSQL は、シングルマスタークラスターとマルチマスタークラスターの両方でサポートされています。マルチマスタークラスターでは、PrestoSQL または PrestoDB を実行するには、外部 Hive メタストアが必要です。「[複数のプライマリノードを持つ Amazon EMR クラスターでサポートされるアプリケーション](#)」を参照してください。
- Docker を使用した Apache Hadoop および Apache Spark での ECR 自動認証サポート: Spark ユーザーは、Docker Hub および Amazon Elastic Container Registry (Amazon ECR) の Docker イメージを使用して、環境とライブラリの依存関係を定義できます。

[Docker を設定し、Amazon EMR 6.x を使用して Docker で Spark アプリケーションを実行します。](#)

- EMR は Apache Hive ACID トランザクションをサポートしています。Amazon EMR 6.1.0 では、データベースの ACID プロパティに準拠するように Hive ACID トランザクションのサポートが追加されています。この機能を使用すると、Amazon Simple Storage Service (Amazon S3) のデータを使用して Hive マネージドテーブルで INSERT, UPDATE, DELETE, および MERGE の各オペレーションを実行できます。これは、ストリーミングの取り込み、データの再記述、MERGE を使用した一括更新、徐々に変化するディメンションなどのユースケースにとって重要な機能です。設定例とユースケースを含め、詳細については、「[Amazon EMR supports Apache Hive ACID transactions](#)」を参照してください。

#### 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールアップリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。

- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケーリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケーリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Apache Flink は EMR 6.0.0 ではサポートされていませんが、Flink 1.11.0 と EMR 6.1.0 ではサポートされています。これは、Hadoop 3 を公式にサポートする Flink の最初のバージョンです。「[Apache Flink 1.11.0 Release Announcement](#)」を参照してください。
- Ganglia はデフォルトの EMR 6.1.0 パッケージバンドルから削除されました。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark

ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に `ulimit` 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、`instance-controller ulimit` を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから `ulimit` を明示的に設定する

1. `/etc/systemd/system/instance-controller.service` を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して `ulimit` を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの `ulimit` を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
```



```
done
sudo systemctl daemon-reload
```

### ⚠ Important

Amazon EMR 6.1.0 および 6.2.0 には、Hudi の挿入、アップサート、および削除操作すべてに大きな影響を及ぼす可能性があるパフォーマンスの問題が含まれています。Amazon EMR 6.1.0 または 6.2.0 で Hudi を使用する予定がある場合は、AWS サポートに連絡してパッチが適用された Hudi RPM を取得する必要があります。

- `spark.driver.extraJavaOptions` と `spark.executor.extraJavaOptions` を使用してカスタムのガベージコレクション設定を指定すると、ガベージコレクション設定の競合のために、EMR 6.1 でドライバー/エグゼキュターの起動に失敗します。EMR リリース 6.1.0 では、代わりにプロパティ `spark.driver.defaultJavaOptions` と `spark.executor.defaultJavaOptions` を使用して、ドライバーとエグゼキュターのカスタムの Spark ガベージコレクション設定を指定する必要があります。詳細については、「[Apache Spark のランタイム環境](#)」および「[Amazon EMR 6.1.0 での Spark ガベージコレクションの設定](#)」を参照してください。
- Oozie で Pig を (Hue 内で。Hue では、Oozie アクションを使用して Pig スクリプトを実行するため) 使用すると、`native-izo` ライブラリをロードできないというエラーが生成されます。このエラーメッセージは情報であるため、Pig の実行はブロックされません。
- Hudi 同時実行のサポート: 現在、Hudi は単一の Hudi テーブルへの同時書き込みをサポートしていません。さらに、Hudi は新しいライターの起動を許可する前に、進行中のライターによって行われた変更をロールバックします。同時書き込みは、このメカニズムを妨げて、競合状態を引き起こし、データの破損につながる可能性があります。データ処理ワークフローの一部として、いつでも 1 つの Hudi テーブルに対して動作する Hudi ライターが 1 つだけであることを確認してください。Hudi は、同じ Hudi テーブルに対して動作する複数の同時リーダーをサポートしています。
- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Amazon EMR 6.1.0 には、Presto を実行しているクラスターに影響する問題があります。長時間 (日単位) 後に、クラスターが「`su: failed to execute /bin/bash: Resource temporarily unavailable`」(`su: /bin/bash` の実行に失敗しました:リソースを一時的に利用できません) または「`shell request failed on channel 0`」(チャンネル 0 でシェルリクエストが失敗しました) などのエラーをスローする可能性があります。この問題は、内部の Amazon EMR プロセス (InstanceController) が生成するライトウェイトプロセス (LWP) が多すぎるため、最終的に Hadoop ユーザーが `nproc` 制限を超える原因となります。これにより、ユーザーは追加のプロセスを開くことができなくなります。この問題の解決策は、EMR 6.2.0 にアップグレードすることです。

## 6.1.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオー

プンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.14.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	2.0.0	EMR S3Select コネクタ
emrfs	2.42.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.11.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	3.2.1-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.2.5	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.2.5	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.2.5	HBase コマンドラインクライアント。
hbase-rest-server	2.2.5	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.2.5	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	3.1.2-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	3.1.2-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	3.1.2-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-2	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-2	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.2-incubating-amzn-2	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.2-incubating-amzn-2	Presto を Hudi で実行するためのバンドルライブラリ。
hudi-prestosql	0.5.2-incubating-amzn-2	PrestoSQL を Hudi で実行するためのバンドルライブラリ。
hudi-spark	0.5.2-incubating-amzn-2	Spark を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hue-server	4.7.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.6.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64+	MariaDB データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.3.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.0.0-HBase-2.0	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	5.0.0-HBase-2.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.232	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.232	いくつかのクエリを実行するサービス。
presto-client	0.232	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
prestoql-coordinator	338	prestoql-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
prestoql-worker	338	いくつかのクエリを実行するサービス。
prestoql-client	338	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト



コンポーネント	Version	説明
ranger-kms-server	2.0.0	Apache Ranger Key Management System
spark-client	3.0.0-amzn-0	Spark コマンドラインクライアント。
spark-history-server	3.0.0-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	3.0.0-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	3.0.0-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.1.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0-preview1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 6.1.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-6.1.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。

分類	説明
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-env	HDFS 環境の値を変更します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。

分類	説明
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive	Apache Hive 用の Amazon EMR-curated 設定。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。

分類	説明
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。

分類	説明
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。

分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
prestosql-log	Presto の log.properties ファイルの値を変更します。
prestosql-config	Presto の config.properties ファイルの値を変更します。
prestosql-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
prestosql-env	Presto の presto-env.sh ファイルの値を変更します。
prestosql-node	PrestoSQL の node.properties ファイルの値を変更します。
prestosql-connector-blackhole	PrestoSQL の blackhole.properties ファイルの値を変更します。
prestosql-connector-cassandra	PrestoSQL の cassandra.properties ファイルの値を変更します。
prestosql-connector-hive	PrestoSQL の hive.properties ファイルの値を変更します。



分類	説明
prestosql-connector-jmx	PrestoSQL の jmx.properties ファイルの値を変更します。
prestosql-connector-kafka	PrestoSQL の kafka.properties ファイルの値を変更します。
prestosql-connector-localfile	PrestoSQL の localfile.properties ファイルの値を変更します。
prestosql-connector-memory	PrestoSQL の memory.properties ファイルの値を変更します。
prestosql-connector-mongodb	PrestoSQL の mongodb.properties ファイルの値を変更します。
prestosql-connector-mysql	PrestoSQL の mysql.properties ファイルの値を変更します。
prestosql-connector-postgresql	PrestoSQL の postgresql.properties ファイルの値を変更します。
prestosql-connector-raptor	PrestoSQL の raptor.properties ファイルの値を変更します。
prestosql-connector-redis	PrestoSQL の redis.properties ファイルの値を変更します。
prestosql-connector-redshift	PrestoSQL の redshift.properties ファイルの値を変更します。
prestosql-connector-tpch	PrestoSQL の tpch.properties ファイルの値を変更します。
prestosql-connector-tpcds	PrestoSQL の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。

分類	説明
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。

分類	説明
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 6.0.1

### 6.0.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Oozie](#)、[Phoenix](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

### 6.0.1 リリースノート

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

#### 変更、拡張、解決した問題

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。

- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケーリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケーリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPCE ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (`$region` を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。 [Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)

## 6.0.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオー

プンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.6	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.14.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.5.0	EMR S3Select コネクタ
emrfs	2.39.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-0.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-0.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-0.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-0.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-0.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-0.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-0.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-0.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	3.2.1-amzn-0.1	個別のノードでコンテナを管理する YARN サービス。



コンポーネント	Version	説明
hadoop-yarn-resourcemanager	3.2.1-amzn-0.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-0.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.2.3	HBase コマンドラインクライアント。
hbase-rest-server	2.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	3.1.2-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	3.1.2-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-0	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-0	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.0-incubating-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.0-incubating-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.0.0	Jupyter ノートブックのマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.5.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64+	MariaDB データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.0.0-HBase-2.0	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.0.0-HBase-2.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.230	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.230	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
presto-client	0.230	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
spark-client	2.4.4	Spark コマンドラインクライアント。
spark-history-server	2.4.4	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.4	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.4	YARN スレーブで必要な Apache Spark ライブラリ。
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0-SNAPSHOT	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 6.0.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-6.0.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。

分類	説明
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-env	HDFS 環境の値を変更します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。

分類	説明
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive	Apache Hive 用の Amazon EMR-curated 設定。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。

分類	説明
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。



分類	説明
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。

分類	説明
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServer の log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。

分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 6.0.0

### 6.0.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Oozie](#)、[Phoenix](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 6.0.0 リリースノート

次のリリースノートには、Amazon EMR リリース 6.0.0 に関する情報が含まれています。

初回リリース日: 2020 年 3 月 10 日

サポートされているアプリケーション

- AWS SDK for Java バージョン 1.11.711
- Ganglia バージョン 3.7.2
- Hadoop バージョン 3.2.1

- HBase バージョン 2.2.3
- HCatalog バージョン 3.1.2
- Hive バージョン 3.1.2
- Hudi バージョン 0.5.0-incubating
- Hue バージョン 4.4.0
- JupyterHub バージョン 1.0.0
- Livy バージョン 0.6.0
- MXNet バージョン 1.5.1
- Oozie バージョン 5.1.0
- Phoenix バージョン 5.0.0
- Presto バージョン 0.230
- Spark バージョン 2.4.4
- TensorFlow バージョン 1.14.0
- Zeppelin バージョン 0.9.0-SNAPSHOT
- Zookeeper バージョン 3.4.14
- コネクタおよびドライバー: DynamoDB Connector 4.14.0

#### Note

Flink、Sqoop、Pig、および Mahout は、Amazon EMR バージョン 6.0.0 では使用できません。

## 新機能

- YARN Docker ランタイムのサポート - Spark ジョブなどの YARN アプリケーションは、Docker コンテナのコンテキストで実行できるようになりました。これにより、Amazon EMR クラスターにカスタムライブラリをインストールすることなく、Docker イメージの依存関係を簡単に定義できます。詳細については、「[Docker 統合の設定](#)」および「[Amazon EMR 6.0.0 を使用して Docker で Spark アプリケーションを実行する](#)」を参照してください。
- Hive LLAP のサポート - クエリのパフォーマンス向上のため、Hive が LLAP 実行モードをサポートしました。詳細については、「[Hive LLAP の使用](#)」を参照してください。

## 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Amazon Linux
  - Amazon Linux 2 は EMR 6.x リリースシリーズのオペレーティングシステムです。
  - systemd は、Amazon Linux 1 で使用される upstart ではなく、サービス管理に使用されます。
- Java Development Kit (JDK)
  - Corretto JDK 8 は、EMR 6.x リリースシリーズのデフォルトの JDK です。
- Scala
  - Scala 2.12 は、Apache Spark および Apache Livy で使用されます。



- Python 3
  - Python 3 が EMR の Python のデフォルトバージョンになりました。
- YARN ノードラベル
  - Amazon EMR 6.x リリースシリーズ以降では、YARN ノードラベル機能はデフォルトで無効になっています。アプリケーションマスタープロセスは、デフォルトでコアノードとタスクノードの両方で実行できます。次のプロパティを設定することで、YARN ノードラベル機能を有効にできます: `yarn.node-labels.enabled` および `yarn.node-labels.am.default-node-label-expression`。詳細については、「[Understanding Primary, Core, and Task Nodes](#)」を参照してください。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース `emr-5.30.x`、`emr-5.31.0`、`emr-5.32.0`、`emr-6.0.0`、`emr-6.1.0`、および `emr-6.2.0` は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の `ulimit` 設定が低くなります。Amazon EMR リリース `5.30.1`、`5.30.2`、`5.31.1`、`5.32.1`、`6.0.1`、`6.1.1`、`6.2.1`、`5.33.0`、`6.3.0` 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの `ulimit` 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の `ulimit` 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に `ulimit` 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、`instance-controller ulimit` を最大の 65536 ファイルに明示的に設定できます。

### コマンドラインから `ulimit` を明示的に設定する

1. `/etc/systemd/system/instance-controller.service` を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

## 2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- 、SparkR PySpark、Spark-shell などの Spark インタラクティブシェルは、追加のライブラリでの Docker の使用をサポートしていません。SparkR
- Amazon EMR バージョン 6.0.0 で Python 3 を使用するには、PATH を `yarn.nodemanager.env-whitelist` に追加する必要があります。
- AWS Glue Data Catalog を Hive のメタストアとして使用する場合、Live Long and Process (LLAP) 機能はサポートされません。
- Spark と Docker の統合で Amazon EMR 6.0.0 を使用する場合、Docker ランタイムで Spark ジョブを送信する際の失敗を回避するために、同じインスタンスタイプおよび同量の EBS ボリュームでクラスター内のインスタンスを設定する必要があります。

- Amazon EMR 6.0.0 では、HBase on Amazon S3 ストレージモードは [HBASE-24286](#) の問題の影響を受けます。既存の S3 データを使用してクラスターを作成する場合、HBase マスターは初期化できません。
- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

#### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 6.0.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.6	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	3.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.14.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.5.0	EMR S3Select コネクタ
emrfs	2.39.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計

コンポーネント	Version	説明
		する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	3.2.1-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	3.2.1-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	3.2.1-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	3.2.1-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	3.2.1-amzn-0	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	3.2.1-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	3.2.1-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	3.2.1-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	3.2.1-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	3.2.1-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	3.2.1-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	2.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	2.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	2.2.3	HBase コマンドラインクライアント。
hbase-rest-server	2.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	2.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	3.1.2-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	3.1.2-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	3.1.2-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	3.1.2-amzn-0	Hive コマンドラインクライアント。
hive-hbase	3.1.2-amzn-0	Hive-hbase クライアント。
hive-metastore-server	3.1.2-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	3.1.2-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.0-incubating-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.0-incubating-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.0.0	Jupyter ノートブックのマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mxnet	1.5.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64+	MariaDB データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	5.0.0-HBase-2.0	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	5.0.0-HBase-2.0	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.230	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.230	いくつかのクエリを実行するサービス。



コンポーネント	Version	説明
presto-client	0.230	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
spark-client	2.4.4	Spark コマンドラインクライアント。
spark-history-server	2.4.4	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.4	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.4	YARN スレーブで必要な Apache Spark ライブラリ。
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.41+	Apache HTTP サーバー。
zeppelin-server	0.9.0-SNAPSHOT	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 6.0.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-6.0.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。

分類	説明
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-env	HDFS 環境の値を変更します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。

分類	説明
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive	Apache Hive 用の Amazon EMR-curated 設定。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。

分類	説明
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。

分類	説明
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。

分類	説明
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServer の log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。



分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR 5.x リリースバージョン

このセクションでは、各 Amazon EMR 5.x リリースバージョンのアプリケーションバージョン、リリースノート、コンポーネントバージョン、および設定分類を示しています。

クラスターを起動するときに、複数の Amazon EMR リリースから選択できます。これにより、互換性の要件に適合するアプリケーションバージョンをテストして使用できます。リリース番号はリリースラベルで指定します。リリースラベルの形式は `emr-x.x.x` です。例えば `emr-7.1.0` です。

新しい Amazon EMR リリースは、特定のリージョンで最初に利用可能になり、その後数日で、他のリージョンでも利用できるようになります。この期間中、お客様のリージョンで最新のリリースバージョンが利用できない場合があります。

すべての Amazon EMR 5.x リリースのアプリケーションバージョンの包括的な表については、「[Amazon EMR 5.x リリースのアプリケーションバージョン](#)」を参照してください。

### トピック

- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR リリース 5.36.2](#)
- [Amazon EMR リリース 5.36.1](#)

- [Amazon EMR リリース 5.36.0](#)
- [Amazon EMR リリース 5.35.0](#)
- [Amazon EMR リリース 5.34.0](#)
- [Amazon EMR リリース 5.33.1](#)
- [Amazon EMR リリース 5.33.0](#)
- [Amazon EMR リリース 5.32.1](#)
- [Amazon EMR リリース 5.32.0](#)
- [Amazon EMR リリース 5.31.1](#)
- [Amazon EMR リリース 5.31.0](#)
- [Amazon EMR リリース 5.30.2](#)
- [Amazon EMR リリース 5.30.1](#)
- [Amazon EMR リリース 5.30.0](#)
- [Amazon EMR リリース 5.29.0](#)
- [Amazon EMR リリース 5.28.1](#)
- [Amazon EMR リリース 5.28.0](#)
- [Amazon EMR リリース 5.27.1](#)
- [Amazon EMR リリース 5.27.0](#)
- [Amazon EMR リリース 5.26.0](#)
- [Amazon EMR リリース 5.25.0](#)
- [Amazon EMR リリース 5.24.1](#)
- [Amazon EMR リリース 5.24.0](#)
- [Amazon EMR リリース 5.23.1](#)
- [Amazon EMR リリース 5.23.0](#)
- [Amazon EMR リリース 5.22.0](#)
- [Amazon EMR リリース 5.21.2](#)
- [Amazon EMR リリース 5.21.1](#)
- [Amazon EMR リリース 5.21.0](#)
- [Amazon EMR リリース 5.20.1](#)

- [Amazon EMR リリース 5.20.0](#)
- [Amazon EMR リリース 5.19.1](#)
- [Amazon EMR リリース 5.19.0](#)
- [Amazon EMR リリース 5.18.1](#)
- [Amazon EMR リリース 5.18.0](#)
- [Amazon EMR リリース 5.17.2](#)
- [Amazon EMR リリース 5.17.1](#)
- [Amazon EMR リリース 5.17.0](#)
- [Amazon EMR リリース 5.16.1](#)
- [Amazon EMR リリース 5.16.0](#)
- [Amazon EMR リリース 5.15.1](#)
- [Amazon EMR リリース 5.15.0](#)
- [Amazon EMR リリース 5.14.2](#)
- [Amazon EMR リリース 5.14.1](#)
- [Amazon EMR リリース 5.14.0](#)
- [Amazon EMR リリース 5.13.1](#)
- [Amazon EMR リリース 5.13.0](#)
- [Amazon EMR リリース 5.12.3](#)
- [Amazon EMR リリース 5.12.2](#)
- [Amazon EMR リリース 5.12.1](#)
- [Amazon EMR リリース 5.12.0](#)
- [Amazon EMR リリース 5.11.4](#)
- [Amazon EMR リリース 5.11.3](#)
- [Amazon EMR リリース 5.11.2](#)
- [Amazon EMR リリース 5.11.1](#)
- [Amazon EMR リリース 5.11.0](#)
- [Amazon EMR リリース 5.10.1](#)
- [Amazon EMR リリース 5.10.0](#)
- [Amazon EMR リリース 5.9.1](#)

- [Amazon EMR リリース 5.9.0](#)
- [Amazon EMR リリース 5.8.3](#)
- [Amazon EMR リリース 5.8.2](#)
- [Amazon EMR リリース 5.8.1](#)
- [Amazon EMR リリース 5.8.0](#)
- [Amazon EMR リリース 5.7.1](#)
- [Amazon EMR リリース 5.7.0](#)
- [Amazon EMR リリース 5.6.1](#)
- [Amazon EMR リリース 5.6.0](#)
- [Amazon EMR リリース 5.5.4](#)
- [Amazon EMR リリース 5.5.3](#)
- [Amazon EMR リリース 5.5.2](#)
- [Amazon EMR リリース 5.5.1](#)
- [Amazon EMR リリース 5.5.0](#)
- [Amazon EMR リリース 5.4.1](#)
- [Amazon EMR リリース 5.4.0](#)
- [Amazon EMR リリース 5.3.2](#)
- [Amazon EMR リリース 5.3.1](#)
- [Amazon EMR リリース 5.3.0](#)
- [Amazon EMR リリース 5.2.3](#)
- [Amazon EMR リリース 5.2.2](#)
- [Amazon EMR リリース 5.2.1](#)
- [Amazon EMR リリース 5.2.0](#)
- [Amazon EMR リリース 5.1.1](#)
- [Amazon EMR リリース 5.1.0](#)
- [Amazon EMR リリース 5.0.3](#)
- [Amazon EMR リリース 5.0.2](#)
- [Amazon EMR リリース 5.0.1](#)
- [Amazon EMR リリース 5.0.0](#)

## Amazon EMR 5.x リリースのアプリケーションバージョン

各 Amazon EMR 5.x リリースで利用可能なアプリケーションバージョンを示した包括的な表については、ブラウザで「[Amazon EMR 5.x リリースのアプリケーションバージョン](#)」を開いて確認してください。

### Amazon EMR リリース 5.36.2

#### 5.36.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Delta](#)、[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
AWS SDK for Java	1.12.206	1.12.206	1.12.206	1.12.159
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.9	2.3.9
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.9
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.267	0.267	0.267	0.266
Spark	2.4.8	2.4.8	2.4.8	2.4.8

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.36.2 リリースノート

次のリリースノートには、Amazon EMR リリース 5.36.2 に関する情報が含まれています。5.36.1 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

### 変更点、機能強化、解決した問題

- このリリースでは、Amazon EMR がクラスターの HDFS レプリケーション係数設定を下回るように、クラスターのスケールダウンロジックが改善されています。この改善により、データの冗長性要件が満たされ、スケーリングオペレーションが停止する可能性が低くなります。
- このリリースでは、Presto または Trino を実行する のクラスタースケールアップワークフローに新しい再試行メカニズムが追加されました。この改善により、1 回のサイズ変更操作が失敗するため、クラスターのサイズ変更が無期限に実行されるリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。
- Amazon EMR がコアノードを正常に廃止し、完全に廃止される前に異常になる間にクラスターのスケールダウンオペレーションが停止する問題を修正しました。
- Amazon EMR が 1 つのノードを再起動するときに、複数のプライマリノードを持つ高可用性クラスター内のノードの安定性が向上します。
- Amazon EC2 で実行されている Amazon EMR でログ管理を最適化します。その結果、クラスターログのストレージコストがわずかに削減される可能性があります。

- プライマリノードで管理される ZooKeeper トランザクションログファイルの管理を改善し、ログファイルが範囲外になり、クラスターオペレーションが中断されるシナリオを最小限に抑えます。
- Yarn と通信できないために、複数のプライマリノードを持つ高可用性クラスターが失敗する可能性があるまれなバグを修正しました ResourceManager。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 503.0	4.14.343	2024 年 xxxxxx 月	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、ア



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			ジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

### 5.36.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK

コンポーネント	Version	説明
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.16.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.5.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.21.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.7.0	EMR S3Select コネクタ
emrfs	2.51.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.14.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.14.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.1-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.1-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.10.1-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.1-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.10.1-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.10.1-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.9-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.9-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.9-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.9-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.9-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.9-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.9-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.10.1-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.10.1-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.10.1-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.13.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.8.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68	MySQL データベースサーバー。
nvidia-cuda	11.0.194	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.267-amzn-1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.267-amzn-1	いくつかのクエリを実行するサービス。
presto-client	0.267-amzn-1	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.8-amzn-2	Spark コマンドラインクライアント。
spark-history-server	2.4.8-amzn-2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.8-amzn-2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.8-amzn-2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.10.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.36.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。



再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-5.36.2 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.

分類	説明	再設定アクション
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

分類	説明	再設定アクション
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.

分類	説明	再設定アクション
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2 .properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.

分類	説明	再設定アクション
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.



分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します	Not available.
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 5.36.2 変更ログ

### 5.36.2 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2024-05-28	デプロイ完了	Amazon EMR 5.36.2 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2024-05-28	ドキュメント公開	Amazon EMR 5.36.2 リリースノートが初めて公開されました
2024-05-20	初回リリース	Amazon EMR 5.36.2 が限定商用リージョンに初めてデプロイされました

## Amazon EMR リリース 5.36.1

### 5.36.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)、[Ju](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
AWS SDK for Java	1.12.206	1.12.206	1.12.159	1.11.970
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.9	2.3.8
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.8
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.267	0.267	0.266	0.261
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.36.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.36.1 に関する情報が含まれています。5.36.0 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

### 変更点、機能強化、解決した問題

- Amazon EMR リリース 5.36.1 では、クラスターのスケールダウン中に Amazon S3 にログをアーカイブするためのサポートが追加されています。以前の 5.x リリースでは、クラスター終了時のみ Amazon S3 にログファイルをアーカイブできました。この改善により、クラスターで生成されたログファイルは、ノードが終了した後も Amazon S3 に残ります。詳細については、「[クラスターのログ記録とデバッグを設定する](#)」を参照してください。

- 5.36.1 リリースでは、クラスター上のログ管理デーモンが改善され、EMR クラスター内の追加のログフォルダを監視できるようになりました。この改善により、ディスクの過剰使用シナリオが最小限に抑えられます。
- 5.36.1 リリースでは、クラスター上のログ管理デーモンが停止すると、自動的に再起動されます。この改善により、ディスクの過剰使用が原因でノードが異常に見えるリスクが軽減されます。
- 5.36.1 リリースでは、プライマリノードの Amazon EMR デーモンが、クラスター内の終了したインスタンスの古いメタデータを保持する問題が修正されています。古いデータを維持すると、クラスター上の CPU とメモリの使用量が際限なく増加し、最終的にはクラスターに障害が発生する可能性があります。
- 複数のプライマリノードで起動されるクラスターの場合、5.36.1 リリースでは、プライマリノードの 1 つで Amazon EC2 ハードウェア障害が発生すると、2 つ目のプライマリノードに障害が発生し、クラスターが不安定になるという問題が修正されています。
- 送信中の暗号化が設定されているクラスターの場合、マネージドスケーリングは Spark シャッフルデータを認識するようになりました。Spark シャッフルデータは、Spark が特定の操作を実行するためにパーティションに再配分するデータです。スケールダウン中、マネージドスケーリングは、シャッフルデータを含むインスタンスを無視します。これにより、料金とパフォーマンスの面でコストがかかる、ジョブの再試行や再計算を防ぐことができます。シャッフル操作の詳細については、[Spark のプログラミングガイド](#)を参照してください。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 4.14.336 223.0		2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)</p>

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 4.14.318 628.0		2023 年 7 月 12 日	連邦)、カナダ (中部)、イスラエル (テルアビブ)  米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

### 5.36.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.16.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-notebook-env</code>	1.5.0	Jupyter エンタープライズゲートウェイを含む <code>emr</code> ノートブック用 Conda env
<code>emr-s3-dist-cp</code>	2.21.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	1.7.0	EMR S3Select コネクタ

コンポーネント	Version	説明
emrfs	2.51.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.14.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.14.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.1-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-4	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.10.1-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.10.1-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.1-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.10.1-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.10.1-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。

コンポーネント	Version	説明
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.9-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.9-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.9-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.9-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.9-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.9-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。



コンポーネント	Version	説明
hive-server2	2.3.9-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.10.1-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.10.1-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.10.1-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.13.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.8.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68	MySQL データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	11.0.194	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.267-amzn-1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.267-amzn-1	いくつかのクエリを実行するサービス。
presto-client	0.267-amzn-1	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。

コンポーネント	Version	説明
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.8-amzn-2	Spark コマンドラインクライアント。
spark-history-server	2.4.8-amzn-2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.8-amzn-2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.8-amzn-2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.10.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.36.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-5.36.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode,

分類	説明	再設定アクション
		Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.

分類	説明	再設定アクション
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.

分類	説明	再設定アクション
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.

分類	説明	再設定アクション
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
hudi-defaults	Hudi の hudi-defaults.conf ファイルの値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.

分類	説明	再設定アクション
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.

分類	説明	再設定アクション
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.

分類	説明	再設定アクション
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j .properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの erver.properties ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの log4j.properties ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します	Not available.
spark-log4j	Spark の log4j.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.

分類	説明	再設定アクション
spark-metrics	Spark の metrics.properties ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。	Not available.
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.

分類	説明	再設定アクション
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## 5.36.1 変更ログ

### 5.36.1 リリースおよびリリースノートの変更ログ

日付	イベント	説明
2023-07-26	更新	新しい OS リリースラベル 2.0.20230612.0 と 2.0.20230628.0 。
2023-05-25	デプロイ完了	Amazon EMR 5.36.1 が <a href="#">サポートされているすべてのリージョンに完全にデプロイされました</a>
2023-05-09	ドキュメント公開	Amazon EMR 5.36.1 リリースノートの初版が公開されました
2023-05-04	初回リリース	Amazon EMR 5.36.1 が限定商用リージョンに初めてデプロイされました

## Amazon EMR リリース 5.36.0

### 5.36.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)、[Ju](#)



以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
AWS SDK for Java	1.12.206	1.12.159	1.11.970	1.11.970
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.8	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.8	2.3.7

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
Hudi	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.267	0.266	0.261	0.245.1
Spark	2.4.8	2.4.8	2.4.8	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.36.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.36.0 に関する情報が含まれています。5.35.0 からの変更が含まれています。

初回リリース日: 2022 年 6 月 15 日

### 新機能

- Amazon EMR リリース 5.36.0 では、Apache Ranger 対応クラスターでの Apache Spark によるデータ定義言語 (DDL) のサポートが追加されています。これにより、Apache Ranger を使用して、Amazon EMR クラスターからのデータベースやテーブルの作成、変更、削除などの操作のためのアクセスを管理できます。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

### Note

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または `ListReleaseLabels` API あるいは `list-release-labels` CLI アクションを使用するときに Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ス

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			トックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 207.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 119.1	4.14.301	2023 年 2 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 004.0	4.14.294	2022 年 11 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 912.1	4.14.291	2022 年 10 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 719.0	4.14.287	2022 年 8 月 10 日	米国西部 (北カリフォルニア)、欧州 (パリ)、欧州 (ストックホルム)、欧州 (フランクフルト)、アジアパシフィック (ムンバイ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 426.0	4.14.281	2022 年 6 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

### 変更、拡張、解決した問題

- Amazon EMR 5.36.0 のアップグレードでは、aws-sdk 1.12.206、Hadoop 2.10.1-amzn-4、Hive 2.3.9-amzn-2、Hudi 0.10.1-amzn-1、Spark 2.4.8-amzn-2、Presto 0.267-amzn-1、Amazon Glue コネクタ 1.18.0、EMRFS 2.51.0 がサポートされるようになりました。



## 既知の問題

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、s3://bucket/table/p=a は s3://bucket/table/p=a b のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の

spark.sql.sources.fastS3PartitionDiscovery.enabled 設定を false にセットします。

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、Log4j2 プロパティのファイル名パターン設定が正しくないため、SecretAgent および RecordServer サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

### 5.36.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.16.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-notebook-env</code>	1.5.0	Jupyter エンタープライズゲートウェイを含む <code>emr</code> ノートブック用 Conda env

コンポーネント	Version	説明
emr-s3-dist-cp	2.21.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.7.0	EMR S3Select コネクタ
emrfs	2.51.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.14.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.14.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。

コンポーネント	Version	説明
hadoop-hdfs-datanode	2.10.1-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.1-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.10.1-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-https-server	2.10.1-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.10.1-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.10.1-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.9-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.9-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.9-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.9-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.9-amzn-2	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	2.3.9-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.9-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.10.1-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.10.1-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.10.1-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.13.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。

コンポーネント	Version	説明
mxnet	1.8.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68	MySQL データベースサーバー。
nvidia-cuda	11.0.194	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.267-amzn-1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.267-amzn-1	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
presto-client	0.267-amzn-1	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.8-amzn-2	Spark コマンドラインクライアント。
spark-history-server	2.4.8-amzn-2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.8-amzn-2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.8-amzn-2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。



コンポーネント	Version	説明
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.10.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.36.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-5.36.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.

分類	説明	再設定アクション
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.

分類	説明	再設定アクション
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

分類	説明	再設定アクション
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.

分類	説明	再設定アクション
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2 .properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.



分類	説明	再設定アクション
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します	Not available.
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 5.35.0

### 5.35.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[Iceberg](#)、[JupyterEnterpriseGateway](#)、[Ju](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
AWS SDK for Java	1.12.159	1.11.970	1.11.970	1.11.970
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
HCatalog	2.3.9	2.3.8	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.8	2.3.7	2.3.7
Hudi	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.266	0.261	0.245.1	0.245.1
Spark	2.4.8	2.4.8	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.35.0 リリースノート

これは Amazon EMR リリース 5.35.0 のリリースノートです。

次のリリースノートには、Amazon EMR リリース 5.35.0 に関する情報が含まれています。5.34.0 からの変更が含まれています。

初回リリース日: 2022 年 3 月 30 日

### 新機能

- Log4j 1.x と Log4j 2.x を使用する Amazon EMR リリース 5.35 アプリケーションは、それぞれ Log4j 1.2.17 (またはそれ以上) と Log4j 2.17.1 (またはそれ以上) を使用するようアップグレードされます。以前のリリースの CVE の問題を軽減するためにブートストラップアクションを使用する必要はなくなります。 [CVE-2021-44228 を軽減するためのアプローチ](#) を参照してください。

### 変更、拡張、解決した問題

#### Flink の変更

タイプの変更	説明
アップグレード	<ul style="list-style-type: none"> <li>Flink のバージョンを 1.14.2 にアップデートします。</li> <li>log4j は 2.17.1 にアップグレードされました。</li> </ul>



## Hadoop の変更

タイプの変更	説明
EMR 5.34.0 以降の Hadoop オープンソース バックポート	<ul style="list-style-type: none"> <li>• <a href="#">YARN-10438</a>: ClientRMService #getContainerReport() で null containerId を処理する</li> <li>• <a href="#">YARN-7266</a>: タイムラインサーバーのイベントハンドラスレッドがロックされている</li> <li>• <a href="#">YARN-10438</a>: ファイルが破損しているか、見つからない場合、ATS 1.5 RollingLevelDb の起動に失敗する</li> <li>• <a href="#">HADOOP-13500</a>: 設定プロパティオブジェクトのイテレーションの同期</li> <li>• <a href="#">YARN-10651</a>: . AbstractYarnScheduler.updateNodeResource() の NPE で CapacityScheduler クラッシュ</li> <li>• <a href="#">HDFS-12221</a>: の xerces を置き換える XmlEditsVisitor</li> <li>• <a href="#">HDFS-16410</a>: での安全でない Xml 解析 OfflineEditsXmlLoader</li> </ul>
Hadoop の変更と修正	<ul style="list-style-type: none"> <li>• KMS と HttpFS で使用されている Tomcat が 8.5.75 にアップグレードされました。</li> <li>• FileSystemOptimizedCommitterV2 では、コミッターの作成中に定義された commitJob 出力パスに成功マーカが書き込まれました。commitJob とタスクレベルの出力パスは異なる場合があるため、マニフェストファイルで定義されているものを使用するようにパスが修正されました。これにより、Hive ジョブで、動的パーティションや UNION ALL などの操作を実行すると、成功マーカが正しく書き込まれるようになりました。</li> </ul>

## Hive の変更点

タイプの変更	説明
<p>Hive がオープンソース <a href="#">リリース 2.3.9</a> にアップグレードされました (これらの JIRA の修正を含む)</p>	<ul style="list-style-type: none"> <li>• HiveConf.java の <a href="#">HIVE-17155</a> : findConfFile() には conf パスにいくつかの問題があります</li> <li>• <a href="#">HIVE-24797</a>: Avro スキーマを解析するときのデフォルト値の検証を無効にする</li> <li>• <a href="#">HIVE-21563</a>: 1 registerAllFunctions回無効にして Table#getEmptyTable performance を改善</li> <li>• <a href="#">HIVE-18147</a>: java.net でテストが失敗する可能性があります。BindException: アドレスは既に使用されています</li> <li>• <a href="#">HIVE-24608</a>: Hive 2.3.x の HMS クライアントで get_table に切り替える</li> <li>• <a href="#">HIVE-21200</a>: ベクトル化 - Parquet UnsupportedOperationException 用の java.lang. をスローする日付列</li> <li>• <a href="#">HIVE-19228</a>: commons-httpclient 3.x の使用を削除する</li> </ul>
<p>EMR 5.34.0 以降の Hive オープンソースバックポート</p>	<ul style="list-style-type: none"> <li>• <a href="#">HIVE-19990</a>: 結合条件に間隔のリテラルを含むクエリが失敗する</li> <li>• <a href="#">HIVE-25824</a>: branch-2.3 を log4j 2.17.0 にアップグレードする</li> <li>• <a href="#">TEZ-4062</a>: 投機的試行のスケジューリングはタスクが完了したら中止する必要がある</li> <li>• <a href="#">TEZ-4108</a>: 投機的実行の競合状態中の NullPointerException</li> <li>• <a href="#">TEZ-3918</a>: tez.task.log.level を設定しても機能しない</li> </ul>
<p>Hive のアップグレードと修正</p>	<ul style="list-style-type: none"> <li>• Log4j バージョンの 2.17.1 へのアップグレード</li> </ul>

タイプの変更	説明
	<ul style="list-style-type: none"> <li>• ORC バージョンの 1.4.3 へのアップグレード</li> <li>• のペナルティスレッドによる FixEDデッドロックを修正しました ShuffleScheduler</li> </ul>
新機能	<ul style="list-style-type: none"> <li>• Hive クエリを AM ログに出力する機能が追加されました。これはデフォルトでは無効になっています。フラグ/設定: <code>tez.am.emr.print.hive.query.in.log</code> 。ステータス (デフォルト): FALSE。</li> </ul>

## Oozie の変更

タイプの変更	説明
EMR 5.34.0 以降の Oozie オープンソースバックポート	<ul style="list-style-type: none"> <li>• <a href="#">OOZIE-3652</a>: Oozie ランチャーは <code>NoSuchFileException</code> 、発生時にディレクトリのリストを再試行する必要があります</li> </ul>

## Pig の変更

タイプの変更	説明
アップグレード	<ul style="list-style-type: none"> <li>• log4j は 1.2.17 にアップグレードされました。</li> </ul>

## 既知の問題

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。

- 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
- 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

### 5.35.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。

コンポーネント	Version	説明
emr-goodies	2.15.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.5.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.20.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.7.0	EMR S3Select コネクタ
emrfs	2.49.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.14.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.14.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.1-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.1-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.10.1-amzn-3	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.1-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.10.1-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.10.1-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.9-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.9-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.9-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.9-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.9-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.9-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.9-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.9.0-amzn-2	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.9.0-amzn-2	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.9.0-amzn-2	Presto を Hudi で実行するためのバンドルライブラリ。



コンポーネント	Version	説明
hue-server	4.10.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.13.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.8.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68	MySQL データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.266-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.266-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.266-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.8-amzn-1	Spark コマンドラインクライアント。
spark-history-server	2.4.8-amzn-1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.8-amzn-1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.8-amzn-1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.10.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.35.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-5.35.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.

分類	説明	再設定アクション
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

分類	説明	再設定アクション
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.



分類	説明	再設定アクション
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2 .properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します	Not available.
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 5.34.0

### 5.34.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
AWS SDK for Java	1.11.970	1.11.970	1.11.970	1.11.890
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13



	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
HCatalog	2.3.8	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.8	2.3.7	2.3.7	2.3.7
Hudi	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.261	0.245.1	0.245.1	0.240.1
Spark	2.4.8	2.4.7	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.9.0	0.9.0	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.34.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.34.0 に関する情報が含まれています。5.33.1 からの変更が含まれています。

初回リリース日: 2022 年 1 月 20 日

更新リリース日: 2022 年 3 月 21 日

### 新機能

- [マネージドスケールリング] Spark シャッフルデータマネージドスケールリング最適化 - Amazon EMR バージョン 5.34.0 以降、および EMR バージョン 6.4.0 以降では、マネージドスケールリングは Spark シャッフルデータ (Spark が特定の操作を実行するためにパーティション間で再分配するデータ) 対応になりました。シャッフル操作の詳細については、「Amazon EMR 管理ガイド」および「[Spark のプログラミングガイド](#)」の「[Amazon EMR での EMR マネージドスケールリングの使用](#)」を参照してください。
- [Hudi] Hudi の設定を簡素化するための改良点。オプティミスティック同時実行制御をデフォルトで無効にしました。

### 変更、拡張、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- 以前は、マルチマスタークラスターでリソースマネージャーを手動で再起動すると、Zookeeper などの Amazon EMR のクラスター上のデーモンが、Zookeeper znode ファイル内の以前に廃止された、または失われたすべてのノードを再ロードしていました。これにより、特定の状況でデフォルトの制限を超えることがありました。Amazon EMR では、1 時間以上経過した、廃止された、

または失われたノードレコードが Zookeeper ファイルから削除されるようになり、内部制限も引き上げられました。

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Zeppelin をバージョン 0.10.0 にアップグレードしました。
- Livy Fix - 0.7.1 にアップグレード
- Spark のパフォーマンスの向上 - EMR 5.34.0 で特定の Spark 設定値がオーバーライドされると、異種工グゼキューターが無効になります。
- WebHDFS と HttpFS サーバーはデフォルトで無効になっています。Hadoop 設定 `dfs.webhdfs.enabled` を使用して WebHDFS を再度有効にすることができます。HttpFS サーバーを起動するには、`sudo systemctl start hadoop-httpfs` を使用します。

## 既知の問題

- Livy ユーザー偽装で使用される Amazon EMR Notebooks 機能は、HttpFS がデフォルトで無効になっているため機能しません。この場合、EMR notebooks は Livy 偽装が有効になっているクラスターに接続できません。回避策は、`sudo systemctl start hadoop-httpfs` を使用して EMR notebooks をクラスターに接続する前に、HttpFS サーバーを起動することです。

- Apache Hadoop HttpFS サーバーはデフォルトで無効になるため、Hue クエリは Amazon EMR 6.4.0 では機能しません。Amazon EMR 6.4.0 で Hue を使用するには、`sudo systemctl start hadoop-httpfs` を使用して Amazon EMR プライマリノードで HttpFS サーバーを手動で起動するか、または [Amazon EMR のステップを使用](#) します。
- Livy ユーザー偽装で使用される Amazon EMR Notebooks 機能は、HttpFS がデフォルトで無効になっているため機能しません。この場合、EMR notebooks は Livy 偽装が有効になっているクラスターに接続できません。回避策は、`sudo systemctl start hadoop-httpfs` を使用して EMR notebooks をクラスターに接続する前に、HttpFS サーバーを起動することです。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## 5.34.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.14.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.4.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.18.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.7.0	EMR S3Select コネクタ
emrfs	2.48.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.13.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。

コンポーネント	Version	説明
flink-jobmanager-config	1.13.1	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.1-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.1-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-hdfs-journalnode	2.10.1-amzn-2	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.1-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.10.1-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.10.1-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.8-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.8-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.8-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.8-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.8-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.8-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.8-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。



コンポーネント	Version	説明
hudi	0.9.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.9.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.9.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.9.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.4.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.1-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.13.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.8.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68	MySQL データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.2.1	Oozie コマンドラインクライアント。
oozie-server	5.2.1	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.261-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.261-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.261-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト

コンポーネント	Version	説明
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.8-amzn-0	Spark コマンドラインクライアント。
spark-history-server	2.4.8-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.8-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.8-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.10.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.34.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-5.34.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ

分類	説明	再設定アクション
		er, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Not available.

分類	説明	再設定アクション
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

分類	説明	再設定アクション
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.



分類	説明	再設定アクション
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2 .properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubs の jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します	Not available.
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 5.33.1

### 5.33.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
AWS SDK for Java	1.11.970	1.11.970	1.11.890	1.11.890
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13



	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.245.1	0.245.1	0.240.1	0.240.1
Spark	2.4.7	2.4.7	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.9.0	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

### 5.33.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.33.0/5.33.1 に関する情報が含まれています。5.32.0 からの変更が含まれています。

初回リリース日: 2021 年 4 月 19 日

最終更新日: 2021 年 8 月 9 日

#### アップグレード

- Amazon Glue コネクタをバージョン 1.15.0 にアップグレードしました
- バージョン 1.11.970 AWS SDK for Java にアップグレード
- EMRFS をバージョン 2.46.0 にアップグレードしました
- EMR Goodies をバージョン 2.14.0 にアップグレードしました
- EMR Record Server をバージョン 1.9.0 にアップグレードしました
- EMR S3 Dist CP をバージョン 2.18.0 にアップグレードしました
- EMR Secret Agent をバージョン 1.8.0 にアップグレードしました
- Flink をバージョン 1.12.1 にアップグレードしました
- Hadoop をバージョン 2.10.1-amzn-1 にアップグレードしました
- Hive をバージョン 2.3.7-amzn-4 にアップグレードしました
- Hudi をバージョン 0.7.0 にアップグレードしました
- Hue をバージョン 4.9.0 にアップグレードしました
- OpenCV をバージョン 4.5.0 にアップグレードしました
- Presto をバージョン 0.245.1-amzn-0 にアップグレードしました
- R をバージョン 4.0.2 にアップグレードしました

- Spark をバージョン 2.4.7-amzn-1 にアップグレードしました
- バージョン 2.4.1 TensorFlow にアップグレード
- Zeppelin をバージョン 0.9.0 にアップグレードしました

#### 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Apache YARN タイムラインサーバーのバージョン 1 および 1.5 のパフォーマンスの問題を修正するためのクラスターの設定

Apache YARN タイムラインサーバーのバージョン 1 および 1.5 では、特に Amazon EMR のデフォルト設定である `yarn.resourcemanager.system-metrics-`

`publisher.enabled=true` を使用する場合に、非常にアクティブで大規模な EMR クラスターでパフォーマンスの問題が発生する可能性があります。オープンソースの YARN タイムラインサーバー v2 では、YARN タイムラインサーバーのスケラビリティに関連するパフォーマンスの問題が解決されています。

この問題の他の回避策には、次のものがあります。

- `yarn-site.xml` での `yarn.resourcemanager.system-metrics-publisher.enabled=false` の設定。
- クラスターの作成時にこの問題の修正を有効にします (以下を参照)。

次の Amazon EMR リリースには、この YARN タイムラインサーバーのパフォーマンス問題の修正が含まれています。

EMR 5.30.2、5.31.1、5.32.1、5.33.1、5.34.x、6.0.1、6.1.1、6.2.1、6.3.1、6.4.x

上記に示されている Amazon EMR リリースで修正を有効にするには、[aws emr create-cluster コマンドのパラメータ](#) `--configurations file://./configurations.json` で渡される設定の JSON ファイルで、以下のプロパティを `true` に設定します。または、[再構成コンソール UI](#) を使用して修正を有効にします。

`configurations.json` ファイルの内容の例

```
[
  {
    "Classification": "yarn-site",
    "Properties": {
      "yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
        "true",
      "yarn.resourcemanager.system-metrics-publisher.enabled": "true"
    },
    "Configurations": []
  }
]
```

- Spark 挿入クエリのために Hive メタストアからパーティションの場所を取得するときに Spark ランタイムが高速になりました。
- コンポーネントのバージョンをアップグレードしました。コンポーネントのバージョンのリストについては、このガイドの「[Amazon EMR リリースについて](#)」を参照してください。

- 新しいクラスターごとに AWS Java SDK バンドルをインストールしました。これは、個々のコンポーネント jar ではなく、すべてのサービス SDK とその依存関係を含む単一の jar です。詳細については、「[Java SDK Bundled Dependency](#)」を参照してください。
- 以前の Amazon EMR リリースでマネージドスケーリングの問題を修正し、アプリケーションの障害率が大幅に低下するように改善しました。
- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPC ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (`$region` を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。[Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)

## 新機能

- Amazon EMR は Amazon S3 Access Points をサポートしています。これは、共有データレイクへのアクセスを簡単に管理できる Amazon S3 の機能です。Amazon S3 アクセスポイントのエイリアスを使用すると、Amazon EMR で大規模なデータアクセスを簡素化できます。Amazon EMR のすべてのバージョンで Amazon S3 アクセスポイントを使用できます。Amazon EMR が利用可能なすべての AWS リージョンでは、追加料金はかかりません。Amazon S3 Access Points と Access Points のエイリアスの詳細については、「Amazon S3 ユーザー ガイド」の「[アクセスポイントにバケット形式のエイリアスの使用](#)」を参照してください。
- Amazon EMR-5.33 は、新しい Amazon EC2 インスタンスタイプ `c5a`、`c5ad`、`c6gn`、`c6gd`、`m6gd`、`d3`、`d3en`、`m5zn`、`r5b`、`r6gd` をサポートしています。「[サポートされるインスタンスタイプ](#)」を参照してください。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース `emr-5.30.x`、`emr-5.31.0`、`emr-5.32.0`、`emr-6.0.0`、`emr-6.1.0`、および `emr-6.2.0` は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の `ulimit` 設定が低くなります。Amazon EMR リリース `5.30.1`、`5.30.2`、`5.31.1`、`5.32.1`、`6.0.1`、`6.1.1`、`6.2.1`、`5.33.0`、`6.3.0` 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの `ulimit` 設定 `4096` になっており、最新の Amazon Linux 2 AMI

の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の `ulimit` 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に `ulimit` 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、`instance-controller ulimit` を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから `ulimit` を明示的に設定する

1. `/etc/systemd/system/instance-controller.service` を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して `ulimit` を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの `ulimit` を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
```

```
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Amazon EMR 6.3.0 および 6.2.0 のプライベートサブネットクラスターで、Ganglia ウェブ UI にアクセスできません。「アクセス拒否 (403)」というエラーが表示されます。Spark、Hue、Zeppelin、Livy JupyterHub、Tez などの他のウェブ UIs は正常に動作しています。パブリックサブネットクラスターでの Ganglia ウェブ UI アクセスも正常に動作します。この問題を解決するには、`sudo systemctl restart httpd` を使用してプライマリノードで httpd サービスを再起動します。この問題は、Amazon EMR 6.4.0 で修正されています。

#### Important

Amazon Linux または Amazon Linux 2 Amazon マシンイメージ (AMI) を実行する EMR クラスターは、デフォルトの Amazon Linux 動作を使用します。再起動が必要な重要かつクリティカルなカーネル更新が自動的にダウンロードされてインストールされることはありません。これは、デフォルトの Amazon Linux AMI を実行している他の Amazon EC2 インスタンスと同じ動作です。Amazon EMR リリースが利用可能になった後に、再起動が必要な新しい Amazon Linux ソフトウェアアップデート (カーネル、NVIDIA、CUDA のアップデートなど) が使用可能になった場合、デフォルトの AMI を実行する EMR クラスターインスタンスで、それらの更新が自動的にダウンロードされてインストールされることはありません。カーネルの更新を取得するには、[Amazon EMR AMI をカスタマイズ](#)して、[最新の Amazon Linux AMI を使用](#)できるようにします。

- AWS Ranger 統合オプションを指定するセキュリティ設定を作成するコンソールのサポートは、現在 GovCloud リージョンではサポートされていません。セキュリティ設定は CLI を使用して実行できます。「Amazon EMR 管理ガイド」の「[EMR セキュリティ設定を作成する](#)」を参照してください。
- スコープ管理ポリシー: AWS ベストプラクティスに合わせて、Amazon EMR では v2 EMR スコープのデフォルト管理ポリシーを廃止されるポリシーの置き換えとして導入しました。「[Amazon EMR 管理ポリシー](#)」を参照してください。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。

- 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
- 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

### 5.33.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。



コンポーネント	Version	説明
emr-goodies	2.14.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.2.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.18.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.6.0	EMR S3Select コネクタ
emrfs	2.46.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.12.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.12.1	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-1.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.1-amzn-1.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-1.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.1-amzn-1.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.10.1-amzn-1.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.1-amzn-1.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-1.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.10.1-amzn-1.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-1.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.10.1-amzn-1.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-1.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.7-amzn-4	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.7-amzn-4	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.7-amzn-4	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.7-amzn-4	Hive コマンドラインクライアント。
hive-hbase	2.3.7-amzn-4	Hive-hbase クライアント。
hive-metastore-server	2.3.7-amzn-4	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.7-amzn-4	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.7.0-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.7.0-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.7.0-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hue-server	4.9.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.2.2	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.7.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MySQL データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.245.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.245.1-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.245.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.7-amzn-1.1	Spark コマンドラインクライアント。
spark-history-server	2.4.7-amzn-1.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.7-amzn-1.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.7-amzn-1.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.9.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.33.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-5.33.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.



分類	説明	再設定アクション
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

分類	説明	再設定アクション
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.

分類	説明	再設定アクション
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2 .properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します	Not available.
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 5.33.0

### 5.33.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
AWS SDK for Java	1.11.970	1.11.890	1.11.890	1.11.852
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.245.1	0.240.1	0.240.1	0.238.3
Spark	2.4.7	2.4.7	2.4.7	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.33.0 リリースノート

### 5.33.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.14.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.2.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.18.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.6.0	EMR S3Select コネクタ
emrfs	2.46.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.12.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.12.1	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.1-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.1-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.10.1-amzn-1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.1-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.10.1-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-1	個別のノードでコンテナを管理する YARN サービス。



コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.10.1-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.7-amzn-4	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.7-amzn-4	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.7-amzn-4	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.7-amzn-4	Hive コマンドラインクライアント。
hive-hbase	2.3.7-amzn-4	Hive-hbase クライアント。
hive-metastore-server	2.3.7-amzn-4	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.7-amzn-4	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.7.0-amzn-1	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.7.0-amzn-1	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.7.0-amzn-1	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.9.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.2.2	Jupyter Notebook のマルチユーザーサーバー

コンポーネント	Version	説明
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.7.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68	MySQL データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.5.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー

コンポーネント	Version	説明
presto-coordinator	0.245.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.245.1-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.245.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	4.0.2	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.7-amzn-1	Spark コマンドラインクライアント。
spark-history-server	2.4.7-amzn-1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.7-amzn-1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.7-amzn-1	YARN スレーブで必要な Apache Spark ライブラリ。

コンポーネント	Version	説明
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.4.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.9.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.33.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-5.33.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-session	Kubernetes/Yarn セッションの Flink log4j-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.



分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

分類	説明	再設定アクション
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.

分類	説明	再設定アクション
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2 .properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します	Not available.
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml の構成設定を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 5.32.1

### 5.32.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
AWS SDK for Java	1.11.890	1.11.890	1.11.852	1.11.852
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.0	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.240.1	0.240.1	0.238.3	0.238.3
Spark	2.4.7	2.4.7	2.4.6	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

### 5.32.1 リリースノート

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

#### 変更、拡張、解決した問題

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース

5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。

- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPC ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (`$region` を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。[Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)

## 既知の問題

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## 5.32.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.13.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.1.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.17.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.6.0	EMR S3Select コネクタ
emrfs	2.45.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.11.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.11.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-0.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.1-amzn-0.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-0.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.1-amzn-0.1	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-hdfs-journalnode	2.10.1-amzn-0.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.1-amzn-0.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-0.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.10.1-amzn-0.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-0.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.10.1-amzn-0.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-0.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。



コンポーネント	Version	説明
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.7-amzn-3	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.7-amzn-3	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.7-amzn-3	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.7-amzn-3	Hive コマンドラインクライアント。
hive-hbase	2.3.7-amzn-3	Hive-hbase クライアント。
hive-metastore-server	2.3.7-amzn-3	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.7-amzn-3	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hudi	0.6.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.6.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.6.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.8.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.7.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68+	MySQL データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.240.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.240.1-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.240.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト

コンポーネント	Version	説明
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.7-amzn-0.1	Spark コマンドラインクライアント。
spark-history-server	2.4.7-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.7-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.7-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.3.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.32.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

#### emr-5.32.1 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the ResourceManager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ

分類	説明	再設定アクション
		er, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegistrator, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。	Not available.

分類	説明	再設定アクション
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.



分類	説明	再設定アクション
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.

分類	説明	再設定アクション
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2 .properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します	Not available.
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 5.32.0

### 5.32.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterEnterpriseGateway](#)、[JupyterHub](#)



以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
AWS SDK for Java	1.11.890	1.11.852	1.11.852	1.11.759
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.6
Hadoop	2.10.1	2.10.0	2.10.0	2.8.5
Hive	2.3.7	2.3.7	2.3.7	2.3.6

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating
Hue	4.8.0	4.7.1	4.7.1	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.240.1	0.238.3	0.238.3	0.232
Spark	2.4.7	2.4.6	2.4.6	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.32.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.32.0 に関する情報が含まれています。5.31.0 からの変更が含まれています。

初回リリース日: 2021 年 1 月 8 日

### アップグレード

- Amazon Glue コネクタをバージョン 1.14.0 にアップグレードしました
- Amazon SageMaker Spark SDK をバージョン 1.4.1 にアップグレードしました
- バージョン 1.11.890 AWS SDK for Java にアップグレード
- EMR DynamoDB Connector バージョン 4.16.0 にアップグレードしました
- EMRFS をバージョン 2.45.0 にアップグレードしました
- EMR Log Analytics Metrics をバージョン 1.18.0 にアップグレードしました
- EMR MetricsAndEventsApiGateway Client をバージョン 1.5.0 にアップグレードしました
- EMR Record Server をバージョン 1.8.0 にアップグレードしました
- EMR S3 Dist CP をバージョン 2.17.0 にアップグレードしました
- EMR Secret Agent をバージョン 1.7.0 にアップグレードしました
- Flink をバージョン 1.11.2 にアップグレードしました
- Hadoop をバージョン 2.10.1-amzn-0 にアップグレードしました
- Hive をバージョン 2.3.7-amzn-3 にアップグレードしました
- Hue をバージョン 4.8.0 にアップグレードしました
- MXNet をバージョン 1.7.0 にアップグレードしました
- OpenCV をバージョン 4.4.0 にアップグレードしました
- Presto をバージョン 0.240.1-amzn-0 にアップグレードしました
- Spark をバージョン 2.4.7-amzn-0 にアップグレードしました
- バージョン 2.3.1 TensorFlow にアップグレード

### 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- コンポーネントのバージョンをアップグレードしました。
- コンポーネントのバージョンのリストについては、このガイドの「[Amazon EMR リリースについて](#)」を参照してください。

## 新機能

- Amazon EMR 5.32.0 および 6.5.0 以降、Apache Spark の動的エグゼキューターサイズ設定はデフォルトで有効になっています。この機能をオンまたはオフにするには、`spark.yarn.heterogeneousExecutors.enabled` 設定パラメータを使用できます。
- インスタンスメタデータサービス (IMDS) V2 サポートステータス: Amazon EMR 5.23.1、5.27.1、および 5.32 以降のコンポーネントは、すべての IMDS 呼び出しに IMDSv2 を使用します。アプリケーションコードでの IMDS 呼び出しでは、IMDSv1 と IMDSv2 の両方を使用するか、または、セ

セキュリティを強化するために IMDSv2 のみを使用するように IMDS を設定できます。その他の 5.x EMR リリースでは、IMDSv1 を無効にすると、クラスターの起動が失敗します。

- Amazon EMR 5.32.0 以降では、Apache Ranger とネイティブに統合されるクラスターを起動できます。Apache Ranger は、Hadoop プラットフォーム全体で包括的なデータセキュリティを有効化、監視、管理するためのオープンソースフレームワークです。詳細については、「[Apache Ranger](#)」を参照してください。ネイティブ統合により、独自の Apache Ranger を導入して、Amazon EMR できめ細かなデータアクセス制御を実施できます。「Amazon EMR 管理ガイド」の「[Amazon EMR と Apache Ranger を統合する](#)」を参照してください。
- Amazon EMR リリース 5.32.0 は Amazon EMR on EKS をサポートしています。EMR on EKS の開始方法の詳細については、「[Amazon EMR on EKS とは](#)」を参照してください。
- Amazon EMR リリース 5.32.0 は Amazon EMR Studio (プレビュー) をサポートしています。EMR Studio の開始方法の詳細については、「[Amazon EMR Studio \(Preview\)](#)」を参照してください。
- スコープ管理ポリシー: AWS ベストプラクティスに合わせて、Amazon EMR では v2 EMR スコープのデフォルト管理ポリシーを廃止されるポリシーの代替として導入しました。「[Amazon EMR 管理ポリシー](#)」を参照してください。

## 既知の問題

- Amazon EMR 6.3.0 および 6.2.0 のプライベートサブネットクラスターで、Ganglia ウェブ UI にアクセスできません。「アクセス拒否 (403)」というエラーが表示されます。Spark、Hue、Zeppelin、Livy JupyterHub、Tez などの他のウェブ UIs は正常に動作しています。パブリックサブネットクラスターでの Ganglia ウェブ UI アクセスも正常に動作します。この問題を解決するには、`sudo systemctl restart httpd` を使用してプライマリノードで httpd サービスを再起動します。この問題は、Amazon EMR 6.4.0 で修正されています。
- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark

ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
```

```
done
sudo systemctl daemon-reload
```

### ⚠ Important

Amazon Linux または Amazon Linux 2 Amazon マシンイメージ (AMI) を実行する EMR クラスターは、デフォルトの Amazon Linux 動作を使用します。再起動が必要な重要かつクリティカルなカーネル更新が自動的にダウンロードされてインストールされることはありません。これは、デフォルトの Amazon Linux AMI を実行している他の Amazon EC2 インスタンスと同じ動作です。Amazon EMR リリースが利用可能になった後に、再起動が必要な新しい Amazon Linux ソフトウェアアップデート (カーネル、NVIDIA、CUDA のアップデートなど) が使用可能になった場合、デフォルトの AMI を実行する EMR クラスターインスタンスで、それらの更新が自動的にダウンロードされてインストールされることはありません。カーネルの更新を取得するには、[Amazon EMR AMI をカスタマイズ](#)して、[最新の Amazon Linux AMI を使用](#)できるようにします。

- Ranger AWS 統合オプションを指定するセキュリティ設定を作成するコンソールのサポートは、現在 GovCloud リージョンではサポートされていません。セキュリティ設定は CLI を使用して実行できます。「Amazon EMR 管理ガイド」の「[EMR セキュリティ設定を作成する](#)」を参照してください。
- Amazon EMR 5.31.0 AtRestEncryption または 5.32.0 を使用するクラスターで または HDFS 暗号化が有効になっている場合、Hive クエリでは次のランタイム例外が発生します。

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
java.lang.RuntimeException: Hive Runtime Error while closing
operators: java.io.IOException: java.util.ServiceConfigurationError:
org.apache.hadoop.security.token.TokenIdentifier: Provider
org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、s3://bucket/table/p=a は s3://bucket/table/p=a b のプレフィックスです。

- 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## 5.32.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.16.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.13.0	Hadoop エコシステムに役立つ追加のライブラリ。



コンポーネント	Version	説明
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.1.0	Jupyter エンタープライズゲートウェイを含む emr ノートブック用 Conda env
emr-s3-dist-cp	2.17.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.6.0	EMR S3Select コネクタ
emrfs	2.45.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.11.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.11.2	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.1-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.1-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.1-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.1-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.10.1-amzn-0	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.1-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.1-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.10.1-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.1-amzn-0	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.10.1-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.1-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.7-amzn-3	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.7-amzn-3	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.7-amzn-3	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.7-amzn-3	Hive コマンドラインクライアント。
hive-hbase	2.3.7-amzn-3	Hive-hbase クライアント。
hive-metastore-server	2.3.7-amzn-3	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.7-amzn-3	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.6.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.6.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.6.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.8.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー

コンポーネント	Version	説明
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.7.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.68	MySQL データベースサーバー。
nvidia-cuda	10.1.243	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー

コンポーネント	Version	説明
presto-coordinator	0.240.1-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.240.1-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.240.1-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.7-amzn-0	Spark コマンドラインクライアント。
spark-history-server	2.4.7-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.7-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.7-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。

コンポーネント	Version	説明
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.3.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.32.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

再設定アクションは、実行中のクラスターのインスタンスグループの設定を指定すると発生します。Amazon EMR によって、変更した分類に対してのみ再設定アクションが開始されます。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## emr-5.32.0 の分類

分類	説明	再設定アクション
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。	Restarts the Resource Manager service.
container-executor	Hadoop YARN の container-executor.cfg ファイルの値を変更します。	Not available.
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。	Not available.
core-site	Hadoop の core-site.xml ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 関連の設定を変更します。	Not available.
emrfs-site	EMRFS の設定を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the



分類	説明	再設定アクション
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml の設定を変更します。	Restarts Flink history server.
flink-log4j	Flink log4j.properties の設定を変更します。	Restarts Flink history server.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。	Not available.
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。	Restarts Flink history server.

分類	説明	再設定アクション
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。	Not available.
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。	Not available.

分類	説明	再設定アクション
hbase	Apache HBase 用の Amazon EMR-curated 設定。	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase の環境の値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。	Not available.
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。	Should not be reconfigured.
hdfs-site	HDFS の hdfs-site.xml の値を変更します。	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

分類	説明	再設定アクション
hcatalog-env	HCatalog 環境の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。	Not available.
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。	Not available.
hive-env	Hive 環境の値を変更します。	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。	Restarts HiveServer2 and HiveMetastore.

分類	説明	再設定アクション
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2 .properties ファイルの値を変更します。	Not available.
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。	Not available.
hive-site	Hive の hive-site.xml ファイルの値を変更します	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します	Not available.
hue-ini	Hue の ini ファイルの値を変更します	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 環境の値を変更します。	Restarts Hadoop Httpfs service.
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。	Not available.
hadoop-kms-env	Hadoop KMS 環境の値を変更します。	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 環境の値を変更します。	Not available.
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。	Not available.
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。	Not available.
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。	Not available.
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。	Not available.
livy-conf	Livy の livy.conf ファイルの値を変更します。	Restarts Livy Server.
livy-env	Livy 環境の値を変更します。	Restarts Livy Server.
livy-log4j	Livy の log4j.properties の設定を変更します。	Restarts Livy Server.
mapred-env	MapReduce アプリケーションの環境の値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie の環境の値を変更します。	Restarts Oozie.

分類	説明	再設定アクション
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。	Restarts Oozie.
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。	Restarts Oozie.
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。	Not available.
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。	Not available.
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。	Not available.
pig-env	Pig 環境の値を変更します。	Not available.
pig-properties	Pig の pig.properties ファイルの値を変更します。	Restarts Oozie.
pig-log4j	Pig の log4j.properties ファイルの値を変更します。	Not available.
presto-log	Presto の log.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-config	Presto の config.properties ファイルの値を変更します。	Restarts Presto-Server.

分類	説明	再設定アクション
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。	Not available.
presto-env	Presto の presto-env.sh ファイルの値を変更します。	Restarts Presto-Server.
presto-node	Presto の node.properties ファイルの値を変更します。	Not available.
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。	Not available.
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。	Not available.
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。	Restarts Presto-Server.
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。	Not available.
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。	Not available.
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。	Not available.
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。	Not available.
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。	Not available.
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。	Not available.



分類	説明	再設定アクション
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。	Not available.
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。	Not available.
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。	Not available.
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。	Not available.
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。	Not available.
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。	Not available.
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 環境の値を変更します。	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。	Not available.
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
recordserver-env	EMR RecordServer 環境の値を変更します。	Restarts EMR record server.
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。	Restarts EMR record server.
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。	Restarts EMR record server.
spark	Apache Spark 用の Amazon EMR-curated 設定。	This property modifies spark-defaults. See actions there.
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-env	Spark 環境の値を変更します。	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します	Not available.
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop の環境の値を変更します。	Not available.
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。	Not available.

分類	説明	再設定アクション
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。	Not available.
tez-site	Tez の tez-site.xml ファイルの値を変更します。	Restarts Oozie.
yarn-env	YARN 環境の値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 環境の値を変更します。	Restarts Zeppelin.
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。	Restarts Zookeeper server.

## Amazon EMR リリース 5.31.1

### 5.31.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[C](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
AWS SDK for Java	1.11.852	1.11.852	1.11.759	1.11.759
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.6	2.3.6
Hadoop	2.10.0	2.10.0	2.8.5	2.8.5
Hive	2.3.7	2.3.7	2.3.6	2.3.6

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.7.1	4.6.0	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.238.3	0.238.3	0.232	0.232
Spark	2.4.6	2.4.6	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

### 5.31.1 リリースノート

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

#### 変更、拡張、解決した問題

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールアップリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPC ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (`$region` を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。 [Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)

## 既知の問題

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

### 5.31.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.4.0	Amazon SageMaker Spark SDK
emr-ddb	4.15.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.13.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.15.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.6.0	EMR S3Select コネクタ
emrfs	2.43.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.11.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.11.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。



コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.0-amzn-0.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.0-amzn-0.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.0-amzn-0.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.0-amzn-0.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.10.0-amzn-0.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.0-amzn-0.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.0-amzn-0.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.10.0-amzn-0.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.0-amzn-0.1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.10.0-amzn-0.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.0-amzn-0.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.7-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.7-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.7-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.7-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.7-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.7-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.7-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.6.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.6.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.6.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。

コンポーネント	Version	説明
hue-server	4.7.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.6.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.3.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.238.3-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.238.3-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.238.3-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.6-amzn-0.1	Spark コマンドラインクライアント。
spark-history-server	2.4.6-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.6-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.6-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.1.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.31.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.31.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。



分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。

分類	説明
jupyter-notebook-conf	Jupyter Notebook の <code>jupyter_notebook_config.py</code> ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの <code>jupyterhub_config.py</code> ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の <code>config.json</code> ファイルの値を変更します。
livy-conf	Livy の <code>livy.conf</code> ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の <code>log4j.properties</code> の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの <code>mapred-site.xml</code> ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の <code>oozie-log4j.properties</code> ファイルの値を変更します。
oozie-site	Oozie の <code>oozie-site.xml</code> ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
phoenix-hbase-site	Phoenix の <code>hbase-site.xml</code> ファイルの値を変更します。

分類	説明
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。

分類	説明
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。
presto-connector-kafka	Presto の <code>kafka.properties</code> ファイルの値を変更します。
presto-connector-localfile	Presto の <code>localfile.properties</code> ファイルの値を変更します。
presto-connector-memory	Presto の <code>memory.properties</code> ファイルの値を変更します。
presto-connector-mongodb	Presto の <code>mongodb.properties</code> ファイルの値を変更します。
presto-connector-mysql	Presto の <code>mysql.properties</code> ファイルの値を変更します。
presto-connector-postgresql	Presto の <code>postgresql.properties</code> ファイルの値を変更します。
presto-connector-raptor	Presto の <code>raptor.properties</code> ファイルの値を変更します。
presto-connector-redis	Presto の <code>redis.properties</code> ファイルの値を変更します。
presto-connector-redshift	Presto の <code>redshift.properties</code> ファイルの値を変更します。
presto-connector-tpch	Presto の <code>tpcj.properties</code> ファイルの値を変更します。
presto-connector-tpcds	Presto の <code>tpcds.properties</code> ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の <code>dbks-site.xml</code> ファイルの値を変更します。

分類	説明
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServerの erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServerの log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.31.0

### 5.31.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[C](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
AWS SDK for Java	1.11.852	1.11.759	1.11.759	1.11.759
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.10.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.6	2.3.6	2.3.6
Hadoop	2.10.0	2.8.5	2.8.5	2.8.5
Hive	2.3.7	2.3.6	2.3.6	2.3.6
Hudi	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.6.0	4.6.0	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.238.3	0.232	0.232	0.232
Spark	2.4.6	2.4.5	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.31.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.31.0 に関する情報が含まれています。5.30.1 からの変更が含まれています。

初回リリース日: 2020 年 10 月 9 日

最終更新日: 2020 年 10 月 15 日



## アップグレード

- Amazon Glue コネクタをバージョン 1.13.0 にアップグレードしました
- Amazon SageMaker Spark SDK をバージョン 1.4.0 にアップグレードしました
- Amazon Kinesis コネクタをバージョン 3.5.9 にアップグレードしました
- を AWS SDK for Java バージョン 1.11.852 にアップグレードしました
- Bigtop-tomcat をバージョン 8.5.56 にアップグレードしました
- EMRFS をバージョン 2.43.0 にアップグレードしました
- EMR MetricsAndEventsApiGateway クライアントをバージョン 1.4.0 にアップグレードしました
- EMR S3 Dist CP をバージョン 2.15.0 にアップグレードしました
- EMR S3 Select をバージョン 1.6.0 にアップグレードしました
- Flink をバージョン 1.11.0 にアップグレードしました
- Hadoop をバージョン 2.10.0 にアップグレードしました
- Hive をバージョン 2.3.7 にアップグレードしました
- Hudi をバージョン 0.6.0 にアップグレードしました
- Hue をバージョン 4.7.1 にアップグレードしました
- JupyterHub をバージョン 1.1.0 にアップグレードしました
- MXNet をバージョン 1.6.0 にアップグレードしました
- OpenCV をバージョン 4.3.0 にアップグレードしました
- Presto をバージョン 0.238.3 にアップグレードしました
- TensorFlow をバージョン 2.1.0 にアップグレードしました

## 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。

- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- [Hive の列統計](#)は、Amazon EMR バージョン 5.31.0 以降でサポートされています。
- コンポーネントのバージョンをアップグレードしました。
- Amazon EMR 5.31.0 での EMRFS S3EC V2 のサポート。S3 Java SDK リリース 1.11.837 以降では、暗号化クライアントバージョン 2 (S3EC V2) がさまざまなセキュリティ機能強化とともに導入されました。詳細については、次を参照してください。
  - S3 ブログ投稿: [Updates to the Amazon S3 encryption client](#)。
  - AWS SDK for Java デベロッパーガイド: [暗号化クライアントと復号クライアントを V2 に移行します](#)。
  - EMR 管理ガイド: [Amazon S3 クライアント側暗号化](#)。

下位互換性を保つため、暗号化クライアント V1 は SDK で引き続き使用できます。

## 新機能

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オーブ

ンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
```

```
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Amazon EMR 5.31.0 を使用すると、Lake Formation と統合されるクラスターを起動できます。この統合により、AWS Glue Data Catalog のデータベースとテーブルに、きめ細かな列レベルのデータフィルタリングが提供されます。また、これにより、企業の ID システムから EMR Notebooks または Apache Zeppelin へのフェデレーションシングルサインオンが可能になります。詳細については、「Amazon EMR 管理ガイド」の「[Amazon EMR と AWS Lake Formation の統合](#)」を参照してください。

Amazon EMR with Lake Formation は現在、米国東部 (オハイオ州とバージニア北部)、米国西部 (北カリフォルニアとオレゴン)、アジアパシフィック (ムンバイ、ソウル、シンガポール、シドニー、東京)、カナダ (中部)、欧州 (フランクフルト、アイルランド、ロンドン、パリ、ストックホルム)、南米 (サンパウロ) の 16 AWS リージョンで利用可能です。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Amazon EMR 5.31.0 AtRestEncryption または 5.32.0 を使用するクラスターで または HDFS 暗号化が有効になっている場合、Hive クエリでは次のランタイム例外が発生します。

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
java.lang.RuntimeException: Hive Runtime Error while closing
operators: java.io.IOException: java.util.ServiceConfigurationError:
org.apache.hadoop.security.token.TokenIdentifier: Provider
org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

### 5.31.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.4.0	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.15.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.13.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.15.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	1.6.0	EMR S3Select コネクタ
<code>emrfs</code>	2.43.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.11.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
flink-jobmanager-config	1.11.0	Apache Flink の EMR ノード上のリソースの管理 JobManager。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.10.0-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.10.0-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.10.0-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.10.0-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-hdfs-journalnode	2.10.0-amzn-0	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.10.0-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.10.0-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.10.0-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.10.0-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resource-manager	2.10.0-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.10.0-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。



コンポーネント	Version	説明
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.7-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.7-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.7-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.7-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.7-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.7-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.7-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hudi	0.6.0-amzn-0	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-spark	0.6.0-amzn-0	Spark を Hudi で実行するためのバンドルライブラリ。
hudi-presto	0.6.0-amzn-0	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.7.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.6.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	4.3.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.238.3-amzn-0	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.238.3-amzn-0	いくつかのクエリを実行するサービス。
presto-client	0.238.3-amzn-0	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト

コンポーネント	Version	説明
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.6-amzn-0	Spark コマンドラインクライアント。
spark-history-server	2.4.6-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.6-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.6-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	2.1.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.31.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.31.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。

分類	説明
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。

分類	説明
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。

分類	説明
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。



分類	説明
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。

分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。

分類	説明
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServer の log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。

分類	説明
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.30.2

### 5.30.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[C](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
AWS SDK for Java	1.11.759	1.11.759	1.11.759	1.11.682
Python	2.7、3.7	2.7、3.7	2.7、3.7	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.10.0	1.9.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.6.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.232	0.227
Spark	2.4.5	2.4.5	2.4.5	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.30.2 リリースノート

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

### 変更、拡張、解決した問題

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールアップリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPC ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (\$region を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。 [Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)

## 既知の問題

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

### 5.30.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。



コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.13.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.14.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.5.0	EMR S3Select コネクタ
emrfs	2.40.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.10.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-6.1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-6.1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-6.1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-6.1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-6.1	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-6.1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-6.1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-6.1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-6.1	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.8.5-amzn-6.1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-6.1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.6-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.6-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.6-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.6-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.6-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.6-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.6-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.2-incubating	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.2-incubating	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.6.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.5.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.232	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.232	いくつかのクエリを実行するサービス。
presto-client	0.232	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.5-amzn-0.1	Spark コマンドラインクライアント。
spark-history-server	2.4.5-amzn-0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.5-amzn-0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.5-amzn-0.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.30.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.30.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。



分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。

分類	説明
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。

分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。

分類	説明
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。

分類	説明
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。
sqoop-site	Sqoop の <code>sqoop-site.xml</code> ファイルの値を変更します。
tez-site	Tez の <code>tez-site.xml</code> ファイルの値を変更します。

分類	説明
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.30.1

### 5.30.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[C](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
AWS SDK for Java	1.11.759	1.11.759	1.11.682	1.11.659
Python	2.7、3.7	2.7、3.7	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.9.1	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1



	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.227	0.227
Spark	2.4.5	2.4.5	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

### 5.30.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.30.1 に関する情報が含まれています。5.30.0 からの変更が含まれています。

初回リリース日: 2020 年 6 月 30 日

最終更新日: 2020 年 8 月 24 日

変更点、機能強化、解決した問題

- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース

- 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- インスタンスコントローラープロセスが無限数のプロセスをスポンする問題を修正しました。
  - Hue が Hive クエリを実行できず、「データベースがロックされています」というメッセージが表示され、クエリの実行が妨げられる問題を修正しました。
  - EMR クラスターでより多くのタスクを同時に実行できるようにするための Spark の問題を修正しました。
  - Jupyter サーバーで「too many files open error」(開いているファイルの数が多すぎるエラー)が発生する、Jupyter Notebook の問題を修正しました。
  - クラスターの開始時間の問題を修正しました。

## 新機能

- Tez UI および YARN タイムラインサーバーの永続アプリケーションインターフェイスは、Amazon EMR バージョン 6.x および EMR バージョン 5.30.1 以降で利用可能です。永続アプリケーション履歴へのワンクリックリンクアクセスにより、SSH 接続でウェブプロキシを設定せずに、ジョブ履歴にすばやくアクセスできます。アクティブなクラスターと終了したクラスターのログは、アプリケーションの終了後 30 日間利用できます。詳細については、「Amazon EMR 管理ガイド」の「[永続アプリケーションユーザーインターフェイスの表示](#)」を参照してください。
- EMR Notebooks 実行 API を使用して、スクリプトまたはコマンドラインで EMR Notebooks を実行できます。AWS コンソールを使用せずに EMR Notebooks の実行を開始、停止、一覧表示、および記述できるので、EMR Notebooks をプログラムで制御できます。パラメータ化されたノートブックセルを使用すると、新しいパラメータ値セットごとにノートブックのコピーを作成しなくても、複数の異なるパラメータ値をノートブックに渡すことができます。「[EMR API のアクション](#)」を参照してください。サンプルコードについては、「[EMR Notebooks をプログラムで実行するサンプルコマンド](#)」を参照してください。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オーブ

ンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
```

```
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

## • EMR Notebooks

クラスターのプライマリノードにカーネルと追加の Python ライブラリをインストールする機能は、EMR バージョン 5.30.1 ではデフォルトで無効になっています。この機能の詳細については、「[クラスターのプライマリノードへのカーネルと Python ライブラリのインストール](#)」を参照してください。

この機能を有効にするには、以下の操作を行います。

1. EMR Notebooks のサービスロールにアタッチされているアクセス許可ポリシーで、次のアクションが許可されていることを確認します。

```
elasticmapreduce:ListSteps
```

詳細については、「[EMR Notebooks のサービスロール](#)」を参照してください。

2. 次の例に示すように、を使用して EMR Notebooks をセットアップするステップをクラスターで AWS CLI 実行します。`us-east-1` を、クラスターが存在するリージョンに置き換えます。詳細については、「[AWS CLIを使用したクラスターへのステップの追加](#)」を参照してください。

```
aws emr add-steps --cluster-id MyClusterID --steps
  Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
```

## • マネージドスケールリング

Presto がインストールされていない 5.30.0 クラスターおよび 5.30.1 クラスターでマネージドスケールリング操作を実行すると、特にスケールダウン操作の後にすぐ、スケールアップ操作が実行されたときに、アプリケーション障害が発生するか、ユニフォームインスタンスグループまたはインスタンスフリートが ARRESTED 状態のままになる場合があります。

回避策として、ジョブに Presto が必要ない場合でも、Amazon EMR リリース 5.30.0 および 5.30.1 を使用してクラスターを作成するときに、インストールするアプリケーションとして Presto を選択します。

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。

- 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

### 5.30.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
emr-goodies	2.13.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.14.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.5.0	EMR S3Select コネクタ
emrfs	2.40.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.10.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。

コンポーネント	Version	説明
hadoop-hdfs-datanode	2.8.5-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-6	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-https-server	2.8.5-amzn-6	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-6	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。



コンポーネント	Version	説明
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.6-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.6-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.6-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.6-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.6-amzn-2	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	2.3.6-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.6-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.2-incubating	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.2-incubating	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.6.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.5.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。

コンポーネント	Version	説明
mariadb-server	5.5.64	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.232	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.232	いくつかのクエリを実行するサービス。
presto-client	0.232	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。

コンポーネント	Version	説明
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.5-amzn-0	Spark コマンドラインクライアント。
spark-history-server	2.4.5-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.5-amzn-0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.5-amzn-0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.30.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.30.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。

分類	説明
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。

分類	説明
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。

分類	説明
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。



分類	説明
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。

分類	説明
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。

分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServer の log4j.properties ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.30.0

### 5.30.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[C](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
AWS SDK for Java	1.11.759	1.11.682	1.11.659	1.11.659
Python	2.7、3.7	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
Delta	-	-	-	-
Flink	1.10.0	1.9.1	1.9.0	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.0.0	1.0.0	1.0.0
Livy	0.7.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.227	0.227	0.227
Spark	2.4.5	2.4.4	2.4.4	2.4.4

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.30.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.30.0 に関する情報が含まれています。5.29.0 からの変更が含まれています。

初回リリース日: 2020 年 5 月 13 日

最終更新日: 2020 年 6 月 25 日

### アップグレード

- バージョン 1.11.759 AWS SDK for Java にアップグレード
- Amazon SageMaker Spark SDK をバージョン 1.3.0 にアップグレードしました
- EMR Record Server をバージョン 1.6.0 にアップグレードしました
- Flink をバージョン 1.10.0 にアップグレードしました
- Ganglia をバージョン 3.7.2 にアップグレードしました
- HBase をバージョン 1.4.13 にアップグレードしました
- Hudi をバージョン 0.5.2 (incubating) にアップグレードしました
- Hue をバージョン 4.6.0 にアップグレードしました
- バージョン 1.1.0 JupyterHub にアップグレード
- Livy をバージョン 0.7.0 (incubating) にアップグレードしました
- Oozie をバージョン 5.2.0 にアップグレードしました

- Presto をバージョン 0.232 にアップグレードしました
- Spark をバージョン 2.4.5 にアップグレードしました
- コネクタとドライバーをアップグレードしました: Amazon Glue Connector 1.12.0、Amazon Kinesis Connector 3.5.0、EMR DynamoDB Connector 4.14.0

## 新機能

- EMR Notebooks - 5.30.0 を使用して作成された EMR クラスターで使用する場合、EMR ノートブックカーネルはクラスター上で実行されます。これにより、ノートブックのパフォーマンスが向上し、カーネルをインストールおよびカスタマイズすることができます。また、クラスターのプライマリノードに Python ライブラリをインストールすることもできます。詳細については、「EMR 管理ガイド」の「[カーネルとライブラリのインストールと使用](#)」を参照してください。
- マネージドスケーリング - Amazon EMR バージョン 5.30.0 以降では、EMR マネージドスケーリングを有効にすることで、ワークロードに応じてクラスター内のインスタンスやユニットの数を自動的に増減できます。Amazon EMR は引き続きクラスターのメトリクスを評価し、クラスターのコストと速度を最適化するためのスケーリングを決定します。詳細については、「Amazon EMR 管理ガイド」の「[クラスターリソースのスケーリング](#)」を参照してください。
- Amazon S3 に保存されているログファイルの暗号化 - Amazon EMR バージョン 5.30.0 以降では、Amazon S3 に保存されているログファイルを AWS KMS カスタマーマネージドキーで暗号化できます。詳細については、「Amazon EMR 管理ガイド」の「[Amazon S3 に保存されているログファイルの暗号化](#)」を参照してください。
- Amazon Linux 2 のサポート - EMR バージョン 5.30.0 以降では、EMR は Amazon Linux 2 OS を使用します。新しいカスタム AMI (Amazon マシンイメージ) は、Amazon Linux 2 AMI に基づいている必要があります。詳細については、「[カスタム AMI の使用](#)」を参照してください。
- Presto Graceful Auto Scale - 5.30.0 を使用する EMR クラスターでは、オートスケーリングのタイムアウト期間を設定することで、Presto タスクの実行が終了するまで待ってからノードの使用を停止できます。詳細については、「[グレースフルな廃止による Presto Auto Scaling の使用](#)」を参照してください。
- 新しい割り当て戦略オプションを使用したフリートインスタンスの作成 — EMR バージョン 5.12.1 以降では、新しい割り当て戦略オプションを使用できます。これにより、クラスターのプロビジョニングにかかる時間が短縮され、スポット割り当てがより正確になり、スポットインスタンスの中断が低減します。デフォルト以外の EMR サービスロールを更新する必要があります。「[インスタンスフリートを構成する](#)」を参照してください。
- `sudo systemctl stop` コマンドと `sudo systemctl start` コマンド — Amazon Linux 2 OS を使用する EMR バージョン 5.30.0 以降では、EMR は `sudo systemctl stop` コマンドおよび `sudo`



systemctl start コマンドを使用してサービスを再起動します。詳細については、「[Amazon EMR でサービスを再起動するにはどうすればよいですか?](#)」を参照してください。

## 変更、機能強化、解決した問題

- EMR バージョン 5.30.0 では、デフォルトで Ganglia がインストールされません。クラスターの作成時に、Ganglia を明示的に選択してインストールできます。
- Spark パフォーマンスの最適化。
- Presto パフォーマンスの最適化。
- Python 3 は、Amazon EMR バージョン 5.30.0 以降ではデフォルトです。
- プライベートサブネット内のサービスアクセス用のデフォルトのマネージドセキュリティグループが更新され、複数の新しいルールが追加されました。サービスアクセスにカスタムセキュリティグループを使用している場合は、同じルールをデフォルトのマネージドセキュリティグループとして含める必要があります。詳細については、「[サービスアクセスの Amazon EMR マネージドセキュリティグループ \(プライベートサブネット\)](#)」を参照してください。Amazon EMR でカスタムサービスロールを使用する場合は、ec2:describeSecurityGroups にアクセス許可を付与して、セキュリティグループが正常に作成されたかどうかを EMR で検証できるようにする必要があります。EMR\_DefaultRole を使用する場合、このアクセス許可はデフォルトのマネージドポリシーに既に含まれています。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- マネージドスケールリング

Presto がインストールされていない 5.30.0 クラスターおよび 5.30.1 クラスターでマネージドスケールリング操作を実行すると、特にスケールダウン操作の後にすぐ、スケールアップ操作が実行されたときに、アプリケーション障害が発生するか、ユニフォームインスタンスグループまたはインスタンスフリートが ARRESTED 状態のままになる場合があります。

回避策として、ジョブに Presto が必要ない場合でも、Amazon EMR リリース 5.30.0 および 5.30.1 を使用してクラスターを作成するときに、インストールするアプリケーションとして Presto を選択します。

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

**Note**

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Hue 4.6.0 のデフォルトのデータベースエンジンは SQLite であり、外部データベースで Hue を使用しようとする、SQLite が原因で問題が発生します。これを修正するには、hue-ini 設定分類

の engine を mysql に設定します。この問題は、Amazon EMR バージョン 5.30.1 で修正されています。

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、s3://bucket/table/p=a は s3://bucket/table/p=a b のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## 5.30.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.13.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.5.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-notebook-env	1.0.0	EMR ノートブック用 Conda env
emr-s3-dist-cp	2.14.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.5.0	EMR S3Select コネクタ
emrfs	2.40.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.10.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-6	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-6	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.8.5-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-6	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.13	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.13	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.13	HBase コマンドラインクライアント。
hbase-rest-server	1.4.13	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.13	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.6-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.6-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.6-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.6-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.6-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.6-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.6-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.2-incubating	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.2-incubating	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.6.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション



コンポーネント	Version	説明
jupyterhub	1.1.0	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.7.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.5.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mariadb-server	5.5.64	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.2.0	Oozie コマンドラインクライアント。
oozie-server	5.2.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.232	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.232	いくつかのクエリを実行するサービス。
presto-client	0.232	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.3	統計的コンピューティング用 R プロジェクト
ranger-kms-server	1.2.0	Apache Ranger Key Management System
spark-client	2.4.5-amzn-0	Spark コマンドラインクライアント。
spark-history-server	2.4.5-amzn-0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.5-amzn-0	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.4.5-amzn-0	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.30.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.30.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。

分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
hudi-env	Hudi 環境の値を変更します。

分類	説明
jupyter-notebook-conf	Jupyter Notebook の <code>jupyter_notebook_config.py</code> ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの <code>jupyterhub_config.py</code> ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の <code>config.json</code> ファイルの値を変更します。
livy-conf	Livy の <code>livy.conf</code> ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の <code>log4j.properties</code> の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの <code>mapred-site.xml</code> ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の <code>oozie-log4j.properties</code> ファイルの値を変更します。
oozie-site	Oozie の <code>oozie-site.xml</code> ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
phoenix-hbase-site	Phoenix の <code>hbase-site.xml</code> ファイルの値を変更します。

分類	説明
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。



分類	説明
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。

分類	説明
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServerの erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServerの log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.29.0

### 5.29.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[C](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
AWS SDK for Java	1.11.682	1.11.659	1.11.659	1.11.615
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.1	1.9.0	1.9.0	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.227	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.29.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.29.0 に関する情報が含まれています。5.28.1 からの変更が含まれています。

初回リリース日: 2020 年 1 月 17 日

## アップグレード

- バージョン 1.11.682 AWS SDK for Java にアップグレード
- Hive をバージョン 2.3.6 にアップグレードしました
- Flink をバージョン 1.9.1 にアップグレードしました
- EMRFS をバージョン 2.38.0 にアップグレードしました
- EMR DynamoDB Connector をバージョン 4.13.0 にアップグレードしました

## 変更、機能強化、解決した問題

- Spark
  - Spark パフォーマンスの最適化。
- EMRFS
  - 管理ガイドでは、整合性のあるビューを実現するために emrfs-site.xml デフォルト設定が更新されています。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.29.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.6	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.13.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.12.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。

コンポーネント	Version	説明
emr-s3-dist-cp	2.13.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.4.0	EMR S3Select コネクタ
emrfs	2.38.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.9.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-5	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-5	ブロックを保存する HDFS ノードレベルのサービス。



コンポーネント	Version	説明
hadoop-hdfs-library	2.8.5-amzn-5	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-5	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-5	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-ftpfs-server	2.8.5-amzn-5	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-5	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-5	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-5	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-5	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-5	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.4.10	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.10	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.10	HBase コマンドラインクライアント。
hbase-rest-server	1.4.10	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.10	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.6-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.6-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.6-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.6-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.6-amzn-1	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	2.3.6-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.6-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.0-incubating	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.0-incubating	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.0.0	Jupyter ノートブックのマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.5.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。

コンポーネント	Version	説明
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.227	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.227	いくつかのクエリを実行するサービス。
presto-client	0.227	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。

コンポーネント	Version	説明
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.4	Spark コマンドラインクライアント。
spark-history-server	2.4.4	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.4	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.4	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 5.29.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.29.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。

分類	説明
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。



分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。

分類	説明
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の musql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。

分類	説明
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServer の log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。

分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.28.1

### 5.28.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[C](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
AWS SDK for Java	1.11.659	1.11.659	1.11.615	1.11.615
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.9.0	1.8.1	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
HCatalog	2.3.6	2.3.6	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.5	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.224	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.28.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.28.1 に関する情報が含まれています。5.28.0 からの変更が含まれています。

初回リリース日: 2020 年 1 月 10 日

### 変更、機能強化、解決した問題

- Spark
  - Spark の互換性の問題が修正されました。
- CloudWatch メトリクス
  - 複数のプライマリノードを持つ EMR クラスターでの Amazon CloudWatch メトリクスの発行を修正しました。
- 無効化されたログメッセージ
  - 「...using old version (<4.5.8) of Apache http client」 (Apache http クライアントの古いバージョン (<4.5.8) を使用) という誤ったログメッセージを無効にしました。

### 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。



### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.28.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.6	Amazon SageMaker Spark SDK

コンポーネント	Version	説明
emr-ddb	4.12.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.11.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.13.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.3.0	EMR S3Select コネクタ
emrfs	2.37.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.9.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコネクタ。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-5	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-5	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-5	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-5	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-5	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-5	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-5	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-5	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-5	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.8.5-amzn-5	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-5	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.10	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.10	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.10	HBase コマンドラインクライアント。
hbase-rest-server	1.4.10	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.10	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.6-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.6-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.6-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.6-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.6-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.6-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.6-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.0-incubating	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.0-incubating	Presto を Hudi で実行するためのバンドルライブラリ。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.0.0	Jupyter ノートブックのマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス

コンポーネント	Version	説明
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.5.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.227	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.227	いくつかのクエリを実行するサービス。
presto-client	0.227	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.4	Spark コマンドラインクライアント。
spark-history-server	2.4.4	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.4	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.4	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。

コンポーネント	Version	説明
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 5.28.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.28.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。



分類	説明
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。

分類	説明
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。

分類	説明
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。

分類	説明
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。

分類	説明
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。

分類	説明
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。

分類	説明
recordserver-log4j	EMR RecordServerの log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.28.0

### 5.28.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hudi](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[C](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
AWS SDK for Java	1.11.659	1.11.615	1.11.615	1.11.595
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12



	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.8.1	1.8.1	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.5	2.3.5	2.3.5
Hudi	0.5.0-incubating	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.2	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
Presto	0.227	0.224	0.224	0.220
Spark	2.4.4	2.4.4	2.4.4	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.28.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.28.0 に関する情報が含まれています。変更は 5.27.0 に関連するものです。

初回リリース日: 2019 年 11 月 12 日

### アップグレード

- Flink をバージョン 1.9.0 にアップグレードしました
- Hive をバージョン 2.3.6 にアップグレードしました
- MXNet をバージョン 1.5.1 にアップグレードしました
- Phoenix をバージョン 4.14.3 にアップグレードしました
- Presto をバージョン 0.227 にアップグレードしました
- Zeppelin をバージョン 0.8.2 にアップグレードしました

## 新機能

- クラスターを作成するときに、Amazon EMR で [Apache Hudi](#) をインストールできるようになりました。詳細については、「[Hudi](#)」を参照してください。
- (2019 年 11 月 25 日) 複数のステップを並行して選択して、クラスター使用率を改善し、コストを削減できるようになりました。また、保留中および実行中のステップの両方をキャンセルできるようになりました。詳細については、「[AWS CLI およびコンソールを使用したステップの操作](#)」を参照してください。
- (2019 年 12 月 3 日) EMR クラスターを で作成して実行できるようになりました AWS Outposts。は、オンプレミス施設でネイティブ AWS サービス、インフラストラクチャ、運用モデル AWS Outposts を有効にします。AWS Outposts 環境では、AWS クラウドで使用するのと同じ AWS APIs ツール、インフラストラクチャを使用できます。詳細については、の「[EMR クラスター AWS Outposts](#)」を参照してください。
- (2020 年 3 月 11 日) Amazon EMR バージョン 5.28.0 以降では、AWS Local Zones をサポートする AWS リージョンの論理拡張として、Local Zones サブネットで Amazon EMR クラスターを作成して実行できます。Local Zone を使用すると、Amazon EMR の機能や、コンピューティングやストレージ AWS サービスなどのサービスのサブセットをユーザーの近くに配置することができ、ローカルで実行されているアプリケーションへのレイテンシーが非常に低くなります。使用可能な Local Zones のリストについては、「[AWS Local Zones](#)」を参照してください。利用可能な AWS ローカルゾーンへのアクセスについては、「[リージョン、アベイラビリティゾーン、およびローカルゾーン](#)」を参照してください。

Local Zones は、現在 Amazon EMR Notebooks をサポートしておらず、インターフェイス VPC エンドポイント (AWS PrivateLink) を使用した Amazon EMR への直接接続をサポートしていません。

## 変更、機能強化、解決した問題

- 高可用性クラスターに対するアプリケーションのサポートを拡大しました。
  - 詳細については、「Amazon EMR 管理ガイド」の「[複数のプライマリノードを持つ Amazon EMR クラスターでサポートされるアプリケーション](#)」を参照してください。
- Spark
  - パフォーマンスの最適化
- [Hive]
  - パフォーマンスの最適化

- Presto
  - パフォーマンスの最適化

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.28.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.6	Amazon SageMaker Spark SDK
emr-ddb	4.12.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.11.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.13.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.3.0	EMR S3Select コネクタ
emrfs	2.37.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.9.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれ

コンポーネント	Version	説明
		た、Hadoop エコシステム アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエー ジェントからメトリクスを集計 する Ganglia メタデータコレ クター。
ganglia-web	3.7.1	Ganglia メタデータコレク ターによって収集されたメトリ クスを表示するウェブアプリ ケーション。
hadoop-client	2.8.5-amzn-5	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクラ イアント。
hadoop-hdfs-datanode	2.8.5-amzn-5	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-5	HDFS コマンドラインライ アントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-5	ファイル名を追跡し、場所を ブロックする HDFS サービ ス。
hadoop-hdfs-journalnode	2.8.5-amzn-5	HA クラスター上の Hadoop ファイルシステムジャーナル を管理するための HDFS サー ビス。
hadoop-httpfs-server	2.8.5-amzn-5	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	2.8.5-amzn-5	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-5	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-5	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-5	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-5	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.10	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.10	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.10	HBase コマンドラインクライアント。
hbase-rest-server	1.4.10	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.10	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	2.3.6-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.6-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.6-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.6-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.6-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.6-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.6-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hudi	0.5.0-incubating	データパイプラインを強化する低レイテンシーかつ高効率な増分処理フレームワーク。
hudi-presto	0.5.0-incubating	Presto を Hudi で実行するためのバンドルライブラリ。



コンポーネント	Version	説明
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.0.0	Jupyter ノートブックのマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.5.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.3-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.3-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.227	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.227	いくつかのクエリを実行するサービス。
presto-client	0.227	Presto サーバーが起動されていない HA クラスターのスタンバイマスターにインストールされる Presto コマンドラインクライアント。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.4	Spark コマンドラインクライアント。
spark-history-server	2.4.4	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.4	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.4.4	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 5.28.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.28.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。

分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。

分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。



分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。

分類	説明
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServerの erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServerの log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。

分類	説明
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.27.1

### 5.27.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
AWS SDK for Java	1.11.615	1.11.615	1.11.595	1.11.566
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.1	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.5
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.2	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.224	0.220	0.220
Spark	2.4.4	2.4.4	2.4.3	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.13.1	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

## 5.27.1 リリースノート

これはパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

インスタンスメタデータサービス (IMDS) V2 サポートステータス: Amazon EMR 5.23.1、5.27.1、および 5.32 以降のコンポーネントは、すべての IMDS 呼び出しに IMDSv2 を使用します。アプリケーションコードでの IMDS 呼び出しでは、IMDSv1 と IMDSv2 の両方を使用するか、または、セキュリティを強化するために IMDSv2 のみを使用するように IMDS を設定できます。その他の 5.x EMR リリースでは、IMDSv1 を無効にすると、クラスターの起動が失敗します。

### 5.27.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.4	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.12.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.11.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.13.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	1.3.0	EMR S3Select コネクタ
<code>emrfs</code>	2.36.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.8.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.10	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.10	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.10	HBase コマンドラインクライアント。



コンポーネント	Version	説明
hbase-rest-server	1.4.10	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.10	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.5-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.5-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.5-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.5-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.5-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.5-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.5-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.0.0	Jupyter ノートブックのマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.4.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.2-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.2-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.224	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.224	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.4	Spark コマンドラインクライアント。
spark-history-server	2.4.4	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.4	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.4	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

### 5.27.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.27.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。

分類	説明
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。



分類	説明
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。

分類	説明
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServerの <code>erver.properties</code> ファイルの値を変更します。
recordserver-log4j	EMR RecordServerの <code>log4j.properties</code> ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の <code>spark-defaults.conf</code> ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の <code>hive-site.xml</code> ファイルの値を変更します
spark-log4j	Spark の <code>log4j.properties</code> ファイルの値を変更します。
spark-metrics	Spark の <code>metrics.properties</code> ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの <code>oraoop-site.xml</code> ファイルの値を変更します。
sqoop-site	Sqoop の <code>sqoop-site.xml</code> ファイルの値を変更します。
tez-site	Tez の <code>tez-site.xml</code> ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。

分類	説明
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.27.0

### 5.27.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
AWS SDK for Java	1.11.615	1.11.595	1.11.566	1.11.546

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
Python	2.7、 3.6	2.7、 3.6	2.7、 3.6	2.7、 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
Phoenix	4.14.2	4.14.2	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.220	0.220	0.219
Spark	2.4.4	2.4.3	2.4.3	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.13.1	1.13.1	1.12.0
Tez	0.9.2	0.9.2	0.9.2	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.13

## 5.27.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.27.0 に関する情報が含まれています。5.26.0 からの変更が含まれています。

初回リリース日: 2019 年 9 月 23 日

### アップグレード

- AWS SDK for Java 1.11.615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- Tensorflow 1.14.0
- コネクタおよびドライバー:
  - DynamoDB Connector 4.12.0

## 新機能

- (2019 年 10 月 24 日) EMR ノートブックの次の新機能は、すべての Amazon EMR リリースで利用できます。
  - Git リポジトリを EMR ノートブックに関連付けて、バージョン管理された環境でノートブックを保存できます。リモートの Git リポジトリを使用して、コードをピアと共有し、既存の Jupyter Notebook を再利用できます。詳細については、「Amazon EMR 管理ガイド」の「[Git リポジトリを Amazon EMR Notebooks に関連付ける](#)」を参照してください。
  - ノートブックの比較とマージを簡素化するために、[nbdime ユーティリティ](#)が EMR ノートブックで利用できるようになりました。
  - EMR ノートブックでサポートされるようになりました JupyterLab。JupyterLab は、Jupyter ノートブックと完全に互換性のあるウェブベースのインタラクティブな開発環境です。これで、ノートブックを JupyterLab または Jupyter Notebook Editor で開くことが選択できるようになりました。
- (2019 年 10 月 30 日) Amazon EMR バージョン 5.25.0 以降では、コンソールでクラスターの [概要] ページまたは [アプリケーションの履歴] タブから Spark 履歴サーバー UI に接続できます。SSH 接続でウェブプロキシを設定する代わりに、Spark 履歴サーバー UI にすばやくアクセスしてアプリケーションのメトリクスを確認し、アクティブなクラスターと終了したクラスターに関連するログファイルにアクセスできます。詳細については、「Amazon EMR 管理ガイド」の「[永続アプリケーションユーザーインターフェイスへのクラスター外アクセス](#)」を参照してください。

## 変更点、機能強化、解決した問題

- 複数のプライマリノードを持つ Amazon EMR クラスター
  - 複数のプライマリノードを持つ Amazon EMR クラスターに Flink をインストールして実行できます。詳細については、「[サポートされるアプリケーションと機能](#)」を参照してください。
  - 複数のプライマリノードを持つ Amazon EMR クラスターに HDFS 透過的暗号化を設定することができます。詳細については、「[HDFS Transparent Encryption on EMR clusters with Multiple Primary Nodes](#)」を参照してください。
  - 複数のプライマリノードを持つ Amazon EMR クラスターで実行されているアプリケーションの設定を変更できるようになりました。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。
- Amazon EMR-DynamoDB Connector

- Amazon EMR-DynamoDB Connector は、boolean、list、map、item、null の DynamoDB データ型をサポートするようになりました。詳細については、「[Hive コマンドを実行するように Hive テーブルをセットアップする](#)」を参照してください。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.27.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通



常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.4	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.12.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.11.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.13.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	1.3.0	EMR S3Select コネクタ
<code>emrfs</code>	2.36.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.8.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。

コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-4	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	2.8.5-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.10	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.10	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.10	HBase コマンドラインクライアント。
hbase-rest-server	1.4.10	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.10	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	2.3.5-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.5-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.5-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.5-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.5-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.5-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.5-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	1.0.0	Jupyter ノートブックのマルチユーザーサーバー

コンポーネント	Version	説明
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.4.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.2-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.2-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー

コンポーネント	Version	説明
presto-coordinator	0.224	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.224	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.4	Spark コマンドラインクライアント。
spark-history-server	2.4.4	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.4	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.4	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.14.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 5.27.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.27.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。

分類	説明
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。



分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。

分類	説明
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。

分類	説明
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。

分類	説明
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。

分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。
recordserver-log4j	EMR RecordServer の log4j.properties ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.26.0

### 5.26.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
AWS SDK for Java	1.11.595	1.11.566	1.11.546	1.11.546
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.9	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.220	0.220	0.219	0.219
Spark	2.4.3	2.4.3	2.4.2	2.4.2



	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.13.1	1.12.0	1.12.0
Tez	0.9.2	0.9.2	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.13	3.4.13

## 5.26.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.26.0 に関する情報が含まれています。5.25.0 からの変更が含まれています。

初回リリース日: 2019 年 8 月 8 日

最終更新日: 2019 年 8 月 19 日

### アップグレード

- AWS SDK for Java 1.11.595
- HBase 1.4.10
- Phoenix 4.14.2
- コネクタおよびドライバー:
  - DynamoDB Connector 4.11.0
  - MariaDB Connector 2.4.2
  - Amazon Redshift JDBC ドライバー 1.2.32.1056

### 新機能

- (ベータ) Amazon EMR 5.26.0 を使用すると、Lake Formation と統合されるクラスターを起動できます。この統合により、AWS Glue Data Catalog のデータベースとテーブルへのきめ細かな列レ

ベルのアクセスが可能になります。また、これにより、企業の ID システムから EMR Notebooks または Apache Zeppelin へのフェデレーションシングルサインオンが可能になります。詳細については、[「Amazon EMR と AWS Lake Formation \( ベータ版\) の統合」](#)を参照してください。

- (2019 年 8 月 19 日) Amazon EMR のパブリックアクセスブロックが、セキュリティグループをサポートしているすべての Amazon EMR リリースで利用できるようになりました。ブロックパブリックアクセスは、各 AWS リージョンに適用されるアカウント全体の設定です。パブリックアクセスブロックにより、クラスターに関連付けられているセキュリティグループに、ポートで IPv4 0.0.0.0/0 または IPv6 :::/0 (パブリックアクセス) からのインバウンドトラフィックを許可するルールがある場合に、クラスターの起動が禁止されます (ポートが例外として指定されている場合を除く)。ポート 22 は、デフォルトで例外になります。詳細については、「Amazon EMR 管理ガイド」の[「Amazon EMR のパブリックアクセスブロックの使用」](#)を参照してください。

## 変更、機能強化、解決した問題

- EMR Notebooks
  - EMR 5.26.0 以降では、EMR Notebooks はデフォルトの Python ライブラリに加えて、ノートブックスコープの Python ライブラリもサポートしています。ノートブックスコープのライブラリは、クラスターを再作成したりノートブックをクラスターに再アタッチしたりすることなく、ノートブックエディター内からインストールできます。ノートブックスコープのライブラリは Python 仮想環境に作成されるため、現在のノートブックセッションにのみ適用されます。これにより、ノートブックの依存関係を分離できます。詳細については、「Amazon EMR 管理ガイド」の[「ノートブックスコープのライブラリの使用」](#)を参照してください。
- EMRFS
  - `fs.s3.consistent.metadata.etag verification.enabled` を true に設定すると、ETag 検証機能 (ベータ) を有効にできます。この機能を有効にすると、EMRFS は Amazon S3 ETags を使用して、読み取られているオブジェクトが利用可能な最新バージョンであることを確認します。この機能は、同じ名前を保持したまま Amazon S3 のファイルが上書きされる read-after-update コースケースに役立ちます。この ETag 検証機能は、現在 S3 Select では使用できません。詳細については、「[整合性のあるビューを設定する](#)」を参照してください。
- Spark
  - 現在、次の最適化がデフォルトで有効になります: ダイナミックパーティションプルーニング、DISTINCT before INTERSECT、DISTINCT クエリが後に続く JOIN の SQL 計画統計推論の改善、スカラーサブクエリのフラット化、結合順序の最適化、ブルームフィルター結合。詳細については、「[Spark のパフォーマンスの最適化](#)」を参照してください。
  - ソートマージ結合の全体的なステージコード生成が改善されました。

- クエリフラグメントとサブクエリの再利用が改善されました。
- Spark の起動時にエグゼキュターを事前割り当てする機能が改善されました。
- 結合の小さい側にブロードキャストヒントが含まれている場合、ブルームフィルター結合は適用されなくなりました。
- Tez
  - Tez に関する問題を解決しました。Tez UI は、複数のプライマリノードを持つ Amazon EMR クラスタで動作するようになりました。

## 既知の問題

- ソートマージ結合の全体的なステージコード生成機能が改善されたことにより、有効にした場合、メモリプレッシャーが増加する可能性があります。この最適化によってパフォーマンスは向上しますが、十分なメモリを提供するように `spark.yarn.executor.memoryOverheadFactor` が調整されていない場合、ジョブの再試行または失敗が生じる可能性があります。この機能を無効にするには、`spark.sql.sortMergeJoinExec.extendedCodegen.enabled` を `false` に設定します。
- 複数のプライマリノードと Kerberos 認証を使用するクラスタの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスタを実行すると、クラスタをしばらく実行した後で、スケールダウンやステップの送信などのクラスタオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスタオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスタのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

**Note**

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.26.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.4	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.11.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.10.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。

コンポーネント	Version	説明
emr-s3-dist-cp	2.12.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.3.0	EMR S3Select コネクタ
emrfs	2.35.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.8.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.8.5-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.4.10	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.10	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.10	HBase コマンドラインクライアント。
hbase-rest-server	1.4.10	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.10	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.5-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.5-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.5-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.5-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.5-amzn-0	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	2.3.5-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.5-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.6	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.4.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット



コンポーネント	Version	説明
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.2-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.2-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.220	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.220	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.3	Spark コマンドラインクライアント。
spark-history-server	2.4.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.13.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 5.26.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.26.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。

分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。

分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。



分類	説明
recordserver-log4j	EMR RecordServerの log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.25.0

### 5.25.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
AWS SDK for Java	1.11.566	1.11.546	1.11.546	1.11.519
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.5	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.4
Livy	0.6.0	0.6.0	0.6.0	0.5.0
MXNet	1.4.0	1.4.0	1.4.0	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
Presto	0.220	0.219	0.219	0.215
Spark	2.4.3	2.4.2	2.4.2	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.12.0	1.12.0	1.12.0
Tez	0.9.2	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.13	3.4.13	3.4.13

## 5.25.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.25.0 に関する情報が含まれています。5.24.1 からの変更が含まれています。

初回リリース日: 2019 年 7 月 17 日

最終更新日: 2019 年 10 月 30 日

### Amazon EMR 5.25.0

#### アップグレード

- AWS SDK for Java 1.11.566
- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- Zookeeper 3.4.14

## 新機能

- (2019 年 10 月 30 日) Amazon EMR バージョン 5.25.0 以降では、コンソールでクラスターの [概要] ページまたは [アプリケーションの履歴] タブから Spark 履歴サーバー UI に接続できます。SSH 接続でウェブプロキシを設定する代わりに、Spark 履歴サーバー UI にすばやくアクセスしてアプリケーションのメトリクスを確認し、アクティブなクラスターと終了したクラスターに関連するログファイルにアクセスできます。詳細については、「Amazon EMR 管理ガイド」の「[永続アプリケーションユーザーインターフェイスへのクラスター外アクセス](#)」を参照してください。

## 変更、機能強化、解決した問題

- Spark
  - ブルームフィルターを使用して入力を事前フィルターすることで、一部の結合のパフォーマンスが改善されました。最適化はデフォルトでは無効になっており、Spark 設定パラメータ `spark.sql.bloomFilterJoin.enabled` を `true` に設定すると、有効にできます。
  - 文字列型の列によるグループ化のパフォーマンスが改善されました。
  - HBase がインストールされていないクラスターの R4 インスタンスタイプのデフォルトの Spark エグゼキューターメモリとコアの設定が改善されました。
  - プルーニングするテーブルが結合の左側にある必要がある、ダイナミックパーティションプルーニング機能に関する以前の問題が解決されました。
  - エイリアスを含む追加のケースに適用できるように、DISTINCT before INTERSECT 最適化が改善されました。
  - DISTINCT クエリが後に続く JOIN の SQL 計画統計推論が改善されました。この改善はデフォルトでは無効になっており、Spark 設定パラメータ `spark.sql.statsImprovements.enabled` を `true` に設定すると、有効にできます。この最適化は Distinct before Intersect 機能で必要であり、`spark.sql.optimizer.distinctBeforeIntersect.enabled` を `true` に設定すると、自動的に有効になります。
  - テーブルのサイズとフィルターに基づいて結合順序が最適化されました。この最適化はデフォルトでは無効になっており、Spark 設定パラメータ `spark.sql.optimizer.sizeBasedJoinReorder.enabled` を `true` に設定すると、有効にできます。

詳細については、「[Spark のパフォーマンスの最適化](#)」を参照してください。

- EMRFS

- EMRFS の設定 `fs.s3.buckets.create.enabled` は、現在、デフォルトでは無効になります。テストでは、この設定を無効にすると、パフォーマンスが向上し、S3 バケットが意図せずに作成されることがなくなることがわかりました。アプリケーションがこの機能に依存する場合は、`emrfs-site` 設定分類でプロパティ `fs.s3.buckets.create.enabled` を `true` に設定すると、この機能を有効にできます。詳細については、「[クラスターの作成時に設定を指定する](#)」を参照してください。
- セキュリティ設定のローカルディスク暗号化と S3 暗号化の改善 (2019 年 8 月 5 日)
  - セキュリティ設定のセットアップで、Amazon S3 暗号化設定をローカルディスク暗号化設定から分離しました。
  - リリース 5.24.0 以降で、EBS 暗号化を有効にするオプションが追加されました。このオプションを選択すると、ストレージボリュームに加えてルートデバイスボリュームが暗号化されます。以前のバージョンでは、ルートデバイスボリュームを暗号化するにはカスタム AMI を使用する必要がありました。
  - 詳細については、「Amazon EMR 管理ガイド」の「[暗号化オプション](#)」を参照してください。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

**Note**

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.25.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.4	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.10.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.9.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。

コンポーネント	Version	説明
emr-s3-dist-cp	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.3.0	EMR S3Select コネクタ
emrfs	2.34.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.8.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。



コンポーネント	Version	説明
hadoop-hdfs-library	2.8.5-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.4.9	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.9	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.9	HBase コマンドラインクライアント。
hbase-rest-server	1.4.9	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.9	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.5-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.5-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.5-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.5-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.5-amzn-0	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	2.3.5-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.5-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.6	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.4.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.1-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.1-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.220	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.220	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.3	Spark コマンドラインクライアント。
spark-history-server	2.4.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.13.1	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.2	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.14	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.14	ZooKeeper コマンドラインクライアント。

## 5.25.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.25.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。

分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。



分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
recordserver-env	EMR RecordServer 環境の値を変更します。
recordserver-conf	EMR RecordServer の erver.properties ファイルの値を変更します。

分類	説明
recordserver-log4j	EMR RecordServerの log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.24.1

### 5.24.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
AWS SDK for Java	1.11.546	1.11.546	1.11.519	1.11.519
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.4	0.9.4
Livy	0.6.0	0.6.0	0.5.0	0.5.0
MXNet	1.4.0	1.4.0	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
Presto	0.219	0.219	0.215	0.215
Spark	2.4.2	2.4.2	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

## 5.24.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.24.1 に関する情報が含まれています。5.24.0 からの変更が含まれています。

初回リリース日: 2019 年 6 月 26 日

変更点、機能強化、解決した問題

- TCP SACK サービス拒否の問題 ([AWS-2019-005](#)) など、重要な Linux カーネルセキュリティ更新を含めるように、Amazon EMR のデフォルトの Amazon Linux AMI を更新しました。

既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエスト

と送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.24.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。



コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.9.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.8.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.3.0	EMR S3Select コネクタ
emrfs	2.33.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.8.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-4	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.8.5-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.9	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.9	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.9	HBase コマンドラインクライアント。
hbase-rest-server	1.4.9	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.9	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.4-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.4-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.4-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.6	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。

コンポーネント	Version	説明
mxnet	1.4.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.1-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.1-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.219	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.219	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。

コンポーネント	Version	説明
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.2	Spark コマンドラインクライアント。
spark-history-server	2.4.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

### 5.24.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.24.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。

分類	説明
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。



分類	説明
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。

分類	説明
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。

分類	説明
phoenix-hbase-metrics	Phoenix の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
phoenix-hbase-site	Phoenix の <code>hbase-site.xml</code> ファイルの値を変更します。
phoenix-log4j	Phoenix の <code>log4j.properties</code> ファイルの値を変更します。
phoenix-metrics	Phoenix の <code>hadoop-metrics2-phoenix.properties</code> ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の <code>pig.properties</code> ファイルの値を変更します。
pig-log4j	Pig の <code>log4j.properties</code> ファイルの値を変更します。
presto-log	Presto の <code>log.properties</code> ファイルの値を変更します。
presto-config	Presto の <code>config.properties</code> ファイルの値を変更します。
presto-password-authenticator	Presto の <code>password-authenticator.properties</code> ファイルの値を変更します。
presto-env	Presto の <code>presto-env.sh</code> ファイルの値を変更します。
presto-node	Presto の <code>node.properties</code> ファイルの値を変更します。
presto-connector-blackhole	Presto の <code>blackhole.properties</code> ファイルの値を変更します。

分類	説明
presto-connector-cassandra	Presto の <code>cassandra.properties</code> ファイルの値を変更します。
presto-connector-hive	Presto の <code>hive.properties</code> ファイルの値を変更します。
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。
presto-connector-kafka	Presto の <code>kafka.properties</code> ファイルの値を変更します。
presto-connector-localfile	Presto の <code>localfile.properties</code> ファイルの値を変更します。
presto-connector-memory	Presto の <code>memory.properties</code> ファイルの値を変更します。
presto-connector-mongodb	Presto の <code>mongodb.properties</code> ファイルの値を変更します。
presto-connector-mysql	Presto の <code>mysql.properties</code> ファイルの値を変更します。
presto-connector-postgresql	Presto の <code>postgresql.properties</code> ファイルの値を変更します。
presto-connector-raptor	Presto の <code>raptor.properties</code> ファイルの値を変更します。
presto-connector-redis	Presto の <code>redis.properties</code> ファイルの値を変更します。
presto-connector-redshift	Presto の <code>redshift.properties</code> ファイルの値を変更します。
presto-connector-tpch	Presto の <code>tpcj.properties</code> ファイルの値を変更します。

分類	説明
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.24.0

### 5.24.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
AWS SDK for Java	1.11.546	1.11.519	1.11.519	1.11.510
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.7.1	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.4	0.9.4	0.9.4
Livy	0.6.0	0.5.0	0.5.0	0.5.0
MXNet	1.4.0	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
Presto	0.219	0.215	0.215	0.215
Spark	2.4.2	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

## 5.24.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.24.0 に関する情報が含まれています。5.23.0 からの変更が含まれています。

初回リリース日: 2019 年 6 月 11 日

最終更新日: 2019 年 8 月 5 日

### アップグレード

- Flink 1.8.0
- Hue 4.4.0
- JupyterHub 0.9.6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK for Java 1.11.546



- コネクタおよびドライバー:
  - DynamoDB Connector 4.9.0
  - MariaDB Connector 2.4.1
  - Amazon Redshift JDBC ドライバー 1.2.27.1051

## 変更、機能強化、解決した問題

- Spark
  - パーティションを動的にプルーニングするための最適化を追加しました。この最適化はデフォルトで無効になっています。これを有効にするには、Spark 設定パラメータ `spark.sql.dynamicPartitionPruning.enabled` を `true` に設定します。
  - INTERSECT のクエリのパフォーマンスが改善されました。この最適化はデフォルトで無効になっています。これを有効にするには、Spark 設定パラメータ `spark.sql.optimizer.distinctBeforeIntersect.enabled` を `true` に設定します。
  - 同じリレーションを使用する集計でスカラーサブクエリをフラット化するための最適化が追加されました。この最適化はデフォルトで無効になっています。これを有効にするには、Spark 設定パラメータ `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` を `true` に設定します。
  - 全体的なステージコード生成が改善されました。

詳細については、「[Spark のパフォーマンスの最適化](#)」を参照してください。

- セキュリティ設定のローカルディスク暗号化と S3 暗号化の改善 (2019 年 8 月 5 日)
  - セキュリティ設定のセットアップで、Amazon S3 暗号化設定をローカルディスク暗号化設定から分離しました。
  - EBS 暗号化を有効にするオプションが追加されました。このオプションを選択すると、ストレージボリュームに加えてルートデバイスボリュームが暗号化されます。以前のバージョンでは、ルートデバイスボリュームを暗号化するにはカスタム AMI を使用する必要がありました。
  - 詳細については、「Amazon EMR 管理ガイド」の「[暗号化オプション](#)」を参照してください。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.24.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオー

プンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.9.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.8.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.3.0	EMR S3Select コネクタ
emrfs	2.33.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.8.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.8.5-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.9	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.9	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.9	HBase コマンドラインクライアント。
hbase-rest-server	1.4.9	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.9	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.4-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.4-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.4-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.4.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.6	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.6.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー

コンポーネント	Version	説明
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.4.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.1-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.1-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.219	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.219	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.2	Spark コマンドラインクライアント。
spark-history-server	2.4.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。



コンポーネント	Version	説明
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

## 5.24.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.24.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。

分類	説明
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。

分類	説明
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。

分類	説明
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。

分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.23.1

### 5.23.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)



## アプリケーションバージョン情報

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
AWS SDK for Java	1.11.519	1.11.519	1.11.510	1.11.479
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.1	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

### 5.23.1 リリースノート

これはパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

インスタンスメタデータサービス (IMDS) V2 サポートステータス: Amazon EMR 5.23.1、5.27.1、および 5.32 以降のコンポーネントは、すべての IMDS 呼び出しに IMDSv2 を使用します。アプリケーションコードでの IMDS 呼び出しでは、IMDSv1 と IMDSv2 の両方を使用するか、または、セキュリティを強化するために IMDSv2 のみを使用するように IMDS を設定できます。その他の 5.x EMR リリースでは、IMDSv1 を無効にすると、クラスターの起動が失敗します。

### 5.23.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.8.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.7.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	1.2.0	EMR S3Select コネクタ
<code>emrfs</code>	2.32.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.7.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-hdfs-journalnode	2.8.5-amzn-3	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.9	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.9	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.9	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.4.9	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.9	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.4-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.4-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.3.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.1-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.1-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.215	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.215	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.0	Spark コマンドラインクライアント。
spark-history-server	2.4.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。



コンポーネント	Version	説明
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

### 5.23.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.23.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。

分類	説明
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。

分類	説明
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.23.0

### 5.23.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。



Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
AWS SDK for Java	1.11.519	1.11.510	1.11.479	1.11.479
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

## 5.23.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.23.0 に関する情報が含まれています。5.22.0 からの変更が含まれています。

初回リリース日: 2019 年 4 月 1 日

最終更新日: 2019 年 4 月 30 日

アップグレード

- AWS SDK for Java 1.11.519

### 新機能

- (2019 年 4 月 30 日) Amazon EMR 5.23.0 以降では、3 つのプライマリノードを持つクラスターを起動して、YARN Resource Manager、HDFS、Spark NameNode、Hive、Ganglia などのアプリケーションの高可用性をサポートできます。プライマリノードは、現在この機能による潜在的な単一障害点ではありません。プライマリノードのいずれかに障害が発生した場合、Amazon EMR は自動的にスタンバイプライマリノードにフェイルオーバーし、障害が発生したプライマリノードを同じ設定とブートストラップアクションを持つ新しいプライマリノードに置き換えます。詳細については、「[プライマリノードの計画と設定](#)」を参照してください。

### 既知の問題

- Tez UI (Amazon EMR リリース 5.26.0 で修正)

Tez UI は、複数のプライマリノードを持つ EMR クラスターでは動作しません。

- Hue (Amazon EMR リリース 5.24.0 で修正)

- Amazon EMR で実行される Hue では、Solr はサポートされません。Amazon EMR リリース 5.20.0 以降、誤設定の問題により Solr が有効になり、次のような無害なエラーメッセージが表示されます。

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Solr のエラーメッセージが表示されないようにするには、以下の操作を行います。

1. SSH を使用してプライマリノードのコマンドラインに接続します。
2. テキストエディタを使用して、hue.ini ファイルを開きます。例:

```
sudo vim /etc/hue/conf/hue.ini
```

3. 語句 `appblacklist` を検索して、行を次のように変更します。

```
appblacklist = search
```

4. 変更を保存して、次の例に示すように Hue を再起動します。

```
sudo stop hue; sudo start hue
```

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

**Note**

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

### 5.23.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.8.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.7.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。

コンポーネント	Version	説明
emr-s3-dist-cp	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.2.0	EMR S3Select コネクタ
emrfs	2.32.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.7.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.8.5-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-hdfs-journalnode	2.8.5-amzn-3	HA クラスター上の Hadoop ファイルシステムジャーナルを管理するための HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.4.9	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.9	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.9	HBase コマンドラインクライアント。
hbase-rest-server	1.4.9	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.9	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.4-amzn-1	Hive-hbase クライアント。



コンポーネント	Version	説明
hive-metastore-server	2.3.4-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.3.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.1	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.1-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.1-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.215	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.215	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.0	Spark コマンドラインクライアント。
spark-history-server	2.4.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

### 5.23.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.23.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。

分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。

分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。



分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。

分類	説明
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.22.0

### 5.22.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
AWS SDK for Java	1.11.510	1.11.479	1.11.479	1.11.479
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.0	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.8	1.4.8	1.4.8

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

## 5.22.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.22.0 に関する情報が含まれています。5.21.0 からの変更が含まれています。

### Important

Amazon EMR リリース 5.22.0 以降、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。以前の Amazon EMR リリースでは、リリースノートで AWS Signature Version 4 のみが表示されていることが示されていない限り、場合によっては Signature Version 2 が使用されます。詳細については、Amazon Simple Storage Service デベロッパーガイドの[「リクエストの認証 \(AWS 署名バージョン 4\)」](#) および [「リクエストの認証 \(AWS 署名バージョン 2\)」](#) を参照してください。

初回リリース日: 2019 年 3 月 20 日

### アップグレード

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- コネクタおよびドライバー:
  - DynamoDB Connector 4.8.0
  - MariaDB Connector 2.2.6

- Amazon Redshift JDBC ドライバー 1.2.20.1043

## 新機能

- EBS 専用ストレージを使用する EC2 インスタンスタイプのデフォルトの EBS 設定を変更しました。Amazon EMR リリース 5.22.0 以降を使用してクラスターを作成する場合、デフォルトの EBS ストレージ容量はインスタンスのサイズに基づいて増加します。さらに、増加したストレージを複数のボリュームに分割することで、IOPS のパフォーマンスが向上しました。別の EBS インスタンスストレージ設定を使用する場合は、EMR クラスターを作成する際、または既存のクラスターをノードに追加する際に指定することができます。インスタンスタイプごとにデフォルトで割り当てられるストレージ容量とボリューム数の詳細については、「Amazon EMR 管理ガイド」の「[インスタンスのデフォルト EBS ストレージ](#)」を参照してください。

## 変更、機能強化、解決した問題

- Spark
  - YARN の Spark 用に新しい設定プロパティ `spark.yarn.executor.memoryOverheadFactor` が導入されました。このプロパティの値は、メモリアーヘッドの値をエグゼキューターメモリのパーセントに設定するスケール係数であり、最小は 384 MB です。`spark.yarn.executor.memoryOverhead` を使用してメモリアーヘッドが明示的に設定されている場合、このプロパティは影響しません。デフォルト値は 0.1875 で、18.75% を表します。Amazon EMR のこのデフォルトは、Spark で内部的に設定されたデフォルトの 10% よりも、エグゼキューターのメモリアーヘッド用に YARN コンテナに余分なスペースを残します。Amazon EMR のデフォルトの 18.75% では、TPC-DS ベンチマークでメモリ関連の障害が減ることが実証されています。
  - パフォーマンスを向上させるために [SPARK-26316](#) をバックポートしました。
- Amazon EMR バージョン 5.19.0、5.20.0、および 5.21.0 では、YARN ノードラベルは HDFS ディレクトリに保存されます。状況によっては、このために、コアノードの起動が遅延し、クラスターのタイムアウトや起動エラーが発生します。Amazon EMR 5.22.0 以降、この問題は解決されています。YARN ノードラベルは各クラスターノードのローカルディスクに保存され、HDFS への依存が回避されています。

## 既知の問題

- Hue (Amazon EMR リリース 5.24.0 で修正)

- Amazon EMR で実行される Hue では、Solr はサポートされません。Amazon EMR リリース 5.20.0 以降、誤設定の問題により Solr が有効になり、次のような無害なエラーメッセージが表示されます。

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Solr のエラーメッセージが表示されないようにするには、以下の操作を行います。

1. SSH を使用してプライマリノードのコマンドラインに接続します。
2. テキストエディタを使用して、hue.ini ファイルを開きます。例:

```
sudo vim /etc/hue/conf/hue.ini
```

3. 語句 `appblacklist` を検索して、行を次のように変更します。

```
appblacklist = search
```

4. 変更を保存して、次の例に示すように Hue を再起動します。

```
sudo stop hue; sudo start hue
```

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.22.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.8.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。



コンポーネント	Version	説明
emr-goodies	2.6.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.2.0	EMR S3Select コネクタ
emrfs	2.31.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.7.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	2.8.5-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.4.9	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.9	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.9	HBase コマンドラインクライアント。
hbase-rest-server	1.4.9	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.9	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.4-amzn-1	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	2.3.4-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.3.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.1.0	Oozie コマンドラインクライアント。
oozie-server	5.1.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.1-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.1-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.215	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.215	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.0	Spark コマンドラインクライアント。
spark-history-server	2.4.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

## 5.22.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.22.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。



分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。

分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。

分類	説明
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.21.2

### 5.21.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
AWS SDK for Java	1.11.479	1.11.479	1.11.479	1.11.461
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.7.0	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

## 5.21.2 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 5.21.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。



コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.5.1	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.2.0	EMR S3Select コネクタ
emrfs	2.30.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.7.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.8	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.8	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.8	HBase コマンドラインクライアント。
hbase-rest-server	1.4.8	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.8	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-hbase	2.3.4-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.4-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.3.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.215	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.215	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.0	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.4.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

## 5.21.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.21.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。

分類	説明
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。



分類	説明
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。

分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.21.1

### 5.21.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
AWS SDK for Java	1.11.479	1.11.479	1.11.461	1.11.461
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.6.2	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.214	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

### 5.21.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.21.1 に関する情報が含まれています。5.21.0 からの変更が含まれています。

初回リリース日: 2019 年 7 月 18 日

#### 変更点、機能強化、解決した問題

- TCP SACK サービス拒否の問題 ([AWS-2019-005](#)) など、重要な Linux カーネルセキュリティ更新を含めるように、Amazon EMR のデフォルトの Amazon Linux AMI を更新しました。

#### 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:



- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

### 5.21.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.1	Amazon SageMaker Spark SDK

コンポーネント	Version	説明
emr-ddb	4.7.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.5.1	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.2.0	EMR S3Select コネクタ
emrfs	2.30.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.7.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.8	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.8	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.8	HBase コマンドラインクライアント。
hbase-rest-server	1.4.8	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.8	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-hbase	2.3.4-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.4-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.3.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.215	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.215	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.0	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.4.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

## 5.21.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.21.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。



分類	説明
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。

分類	説明
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。

分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.21.0

### 5.21.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
AWS SDK for Java	1.11.479	1.11.461	1.11.461	1.11.433
Python	2.7、3.6	2.7、3.6	2.7、3.6	2.7、3.4
Scala	2.11.12	2.11.12	2.11.12	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
Delta	-	-	-	-
Flink	1.7.0	1.6.2	1.6.2	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.7
HCatalog	2.3.4	2.3.4	2.3.4	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.214	0.214	0.212
Spark	2.4.0	2.4.0	2.4.0	2.3.2



	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.11.0
Tez	0.9.1	0.9.1	0.9.1	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

## 5.21.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.21.0 に関する情報が含まれています。5.20.0 からの変更が含まれています。

リリース日: 2019 年 2 月 18 日

最終更新日: 2019 年 4 月 3 日

### アップグレード

- Flink 1.7.0
- Presto 0.215
- AWS SDK for Java 1.11.479

### 新機能

- (2019 年 4 月 3 日) Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## 変更、機能強化、解決した問題

- Zeppelin
  - [ZEPPELIN-3878](#) をバックポートしました。

## 既知の問題

- Hue (Amazon EMR リリース 5.24.0 で修正)
  - Amazon EMR で実行される Hue では、Solr はサポートされません。Amazon EMR リリース 5.20.0 以降、誤設定の問題により Solr が有効になり、次のような無害なエラーメッセージが表示されます。

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Solr のエラーメッセージが表示されないようにするには、以下の操作を行います。

1. SSH を使用してプライマリノードのコマンドラインに接続します。
2. テキストエディタを使用して、hue.ini ファイルを開きます。例:

```
sudo vim /etc/hue/conf/hue.ini
```

3. 語句 appblacklist を検索して、行を次のように変更します。

```
appblacklist = search
```

4. 変更を保存して、次の例に示すように Hue を再起動します。

```
sudo stop hue; sudo start hue
```

- Tez
  - この問題は、Amazon EMR 5.22.0 で修正されています。

クラスタープライマリノードへの SSH 接続を使用して `http://MasterDNS:8080/tez-ui` で Tez UI に接続すると、「Adapter operation failed - Timeline server (ATS) is out of reach. Either it is down, or CORS is not enabled」(~~アダプタの操作に失敗しました - タイムラインサーバ (ATS)~~)

に到達できません。停止しているか、または CORS が有効になっていません) というエラーが表示されるか、タスクで予期せずに N/A と表示されます。

これは、Tez UI がプライマリノードのホスト名ではなく localhost を使用して YARN タイムラインサーバーにリクエストを行ったことが原因で発生します。回避策として、ブートストラップアクションまたはステップとしてスクリプトを実行できます。このスクリプトで、Tez configs.env ファイル内のホスト名を更新します。スクリプトの詳細と場所については、「[Bootstrap Instructions](#)」を参照してください。

- Amazon EMR バージョン 5.19.0、5.20.0、および 5.21.0 では、YARN ノードラベルは HDFS デイレクトリに保存されます。状況によっては、このために、コアノードの起動が遅延し、クラスターのタイムアウトや起動エラーが発生します。Amazon EMR 5.22.0 以降、この問題は解決されています。YARN ノードラベルは各クラスターノードのローカルディスクに保存され、HDFS への依存が回避されています。
- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

**Note**

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.21.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.7.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.5.1	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。

コンポーネント	Version	説明
emr-s3-dist-cp	2.11.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.2.0	EMR S3Select コネクタ
emrfs	2.30.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.7.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.8.5-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.8	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.8	1 つ以上の HBase リージョンに対応するサービス。

コンポーネント	Version	説明
hbase-client	1.4.8	HBase コマンドラインクライアント。
hbase-rest-server	1.4.8	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.8	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.4-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.4-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.3.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。



コンポーネント	Version	説明
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.215	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.215	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.0	Spark コマンドラインクライアント。
spark-history-server	2.4.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

### 5.21.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.21.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。

分類	説明
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。

分類	説明
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。



分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.20.1

### 5.20.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
AWS SDK for Java	1.11.461	1.11.461	1.11.433	1.11.433
Python	2.7、3.6	2.7、3.6	2.7、3.4	2.7、3.4
Scala	2.11.12	2.11.12	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.2	1.6.2	1.6.1	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.7	1.4.7
HCatalog	2.3.4	2.3.4	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.2.0	4.2.0

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.0	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.214	0.212	0.212
Spark	2.4.0	2.4.0	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.11.0	1.11.0
Tez	0.9.1	0.9.1	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

## 5.20.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 5.20.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.2.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.7.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.5.1	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.2.0	EMR S3Select コネクタ
emrfs	2.29.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.6.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。

コンポーネント	Version	説明
hadoop-hdfs-datanode	2.8.5-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.8	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。

コンポーネント	Version	説明
hbase-region-server	1.4.8	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.8	HBase コマンドラインクライアント。
hbase-rest-server	1.4.8	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.8	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.4-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.4-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。

コンポーネント	Version	説明
hive-server2	2.3.4-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.3.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。



コンポーネント	Version	説明
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.214	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.214	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.0	Spark コマンドラインクライアント。
spark-history-server	2.4.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.4.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.0	YARN スレーブで必要な Apache Spark ライブラリ。

コンポーネント	Version	説明
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

### 5.20.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.20.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。

分類	説明
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。

分類	説明
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。

分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。

分類	説明
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。



分類	説明
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.20.0

### 5.20.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
AWS SDK for Java	1.11.461	1.11.433	1.11.433	1.11.393
Python	2.7、3.6	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.12	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.2	1.6.1	1.6.1	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.7	1.4.7	1.4.7
HCatalog	2.3.4	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.4
Hive	2.3.4	2.3.3	2.3.3	2.3.3

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
Hudi	-	-	-	-
Hue	4.3.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.0	1.3.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.212	0.212	0.210
Spark	2.4.0	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.11.0	1.11.0	1.9.0
Tez	0.9.1	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.12

## 5.20.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.20.0 に関する情報が含まれています。5.19.0 からの変更が含まれています。

初回リリース日: 2018 年 12 月 18 日

最終更新: 2019 年 1 月 22 日

### アップグレード

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214
- Spark 2.4.0
- TensorFlow 1.12.0
- Tez 0.9.1
- AWS SDK for Java 1.11.461

### 新機能

- (2019 年 1 月 22 日) Amazon EMR の Kerberos は、外部 KDC からのプリンシパルの認証をサポートするように改善されました。これにより、複数のクラスターが単一の外部 KDC を共有できるため、プリンシパル管理が集中化されます。さらに、外部 KDC は Active Directory ドメインとのクロス領域信頼を得られます。これにより、すべてのクラスターが Active Directory からプリンシパルを認証できます。詳しくは、「Amazon EMR 管理ガイド」の「[Kerberos 認証を使用する](#)」を参照してください。

### 変更、機能強化、解決した問題

- Amazon EMR のデフォルトの Amazon Linux AMI
  - Python3 パッケージが python 3.4 から 3.6 にアップグレードされました。

- EMRFS S3 向けに最適化されたコミッター
  - 現在、EMRFS S3 向けに最適化されたコミッターがデフォルトで有効になるため、書き込みパフォーマンスが向上します。詳細については、「[EMRFS S3 向けに最適化されたコミッターを使用する](#)」を参照してください。
- [Hive]
  - [HIVE-16686](#) をバックポートしました。
- Spark と Hive での Glue
  - EMR 5.20.0 以降では、AWS Glue Data Catalog をメタストアとして使用すると、Spark と Hive の並列パーティションプルーニングが自動的に有効になります。この変更により、複数のリクエストを並列に実行してパーティションを取得できるため、クエリ計画時間が大幅に短縮されます。同時に実行できるセグメントの総数は、1~10 の範囲です。デフォルト値は 5 であり、これが推奨される設定です。hive-site 設定分類の aws.glue.partition.num.segments プロパティを指定すると、これを変更できます。スロットリングが発生した場合は、値を 1 に変更して機能をオフにすることができます。詳細については、「[AWS Glue セグメント構造](#)」を参照してください。

## 既知の問題

- Hue (Amazon EMR リリース 5.24.0 で修正)
  - Amazon EMR で実行される Hue では、Solr はサポートされません。Amazon EMR リリース 5.20.0 以降、誤設定の問題により Solr が有効になり、次のような無害なエラーメッセージが表示されます。

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Solr のエラーメッセージが表示されないようにするには、以下の操作を行います。

1. SSH を使用してプライマリノードのコマンドラインに接続します。
2. テキストエディタを使用して、hue.ini ファイルを開きます。例:

```
sudo vim /etc/hue/conf/hue.ini
```

3. 語句 appblacklist を検索して、行を次のように変更します。

```
appblacklist = search
```

- 変更を保存して、次の例に示すように Hue を再起動します。

```
sudo stop hue; sudo start hue
```

- Tez

- この問題は、Amazon EMR 5.22.0 で修正されています。

クラスタープライマリノードへの SSH 接続を使用して `http://MasterDNS:8080/tez-ui` で Tez UI に接続すると、「Adapter operation failed - Timeline server (ATS) is out of reach. Either it is down, or CORS is not enabled」(アダプタの操作に失敗しました - タイムラインサーバー (ATS) に到達できません。停止しているか、または CORS が有効になっていません) というエラーが表示されるか、タスクで予期せずに N/A と表示されます。

これは、Tez UI がプライマリノードのホスト名ではなく `localhost` を使用して YARN タイムラインサーバーにリクエストを行ったことが原因で発生します。回避策として、ブートストラップアクションまたはステップとしてスクリプトを実行できます。このスクリプトで、`Tez configs.env` ファイル内のホスト名を更新します。スクリプトの詳細と場所については、「[Bootstrap Instructions](#)」を参照してください。

- Amazon EMR バージョン 5.19.0、5.20.0、および 5.21.0 では、YARN ノードラベルは HDFS デイレクトリに保存されます。状況によっては、このために、コアノードの起動が遅延し、クラスターのタイムアウトや起動エラーが発生します。Amazon EMR 5.22.0 以降、この問題は解決されています。YARN ノードラベルは各クラスターノードのローカルディスクに保存され、HDFS への依存が回避されています。
- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。

- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## 5.20.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。

コンポーネント	Version	説明
emr-goodies	2.5.1	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.2.0	EMR S3Select コネクタ
emrfs	2.29.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.6.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。



コンポーネント	Version	説明
hadoop-client	2.8.5-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.4.8	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.8	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.8	HBase コマンドラインクライアント。
hbase-rest-server	1.4.8	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.8	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.4-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.4-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.4-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.4-amzn-0	Hive-hbase クライアント。

コンポーネント	Version	説明
hive-metastore-server	2.3.4-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.4-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.3.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.1	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.214	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.214	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.4.0	Spark コマンドラインクライアント。
spark-history-server	2.4.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.4.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.4.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.12.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.9.1	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

## 5.20.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.20.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。

分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。



分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。

分類	説明
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.19.1

### 5.19.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
AWS SDK for Java	1.11.433	1.11.433	1.11.393	1.11.393
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.1	1.6.0	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.7

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.0	1.3.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.212	0.210	0.210
Spark	2.3.2	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.11.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.12	3.4.12

### 5.19.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.19.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.0	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.5.1	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.1.0	EMR S3Select コネクタ
emrfs	2.28.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.6.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。



コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.5-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.8.5-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.7	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.7	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.7	HBase コマンドラインクライアント。
hbase-rest-server	1.4.7	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.7	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-2	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-hbase	2.3.3-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.212	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.212	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.2	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.3.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

## 5.19.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.19.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。

分類	説明
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。

分類	説明
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。



分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.19.0

### 5.19.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
AWS SDK for Java	1.11.433	1.11.393	1.11.393	1.11.336
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
Delta	-	-	-	-
Flink	1.6.1	1.6.0	1.6.0	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.210	0.210	0.206
Spark	2.3.2	2.3.2	2.3.2	2.3.1

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.7.3
ZooKeeper	3.4.13	3.4.12	3.4.12	3.4.12

## 5.19.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.19.0 に関する情報が含まれています。5.18.0 からの変更が含まれています。

リリース日: 2018 年 11 月 7 日

最終更新日: 2018 年 11 月 19 日

### アップグレード

- Hadoop 2.8.5
- Flink 1.6.1
- JupyterHub 0.9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1.11.0
- Zookeeper 3.4.13
- AWS SDK for Java 1.11.433

## 新機能

- (2018 年 11 月 19 日) EMR Notebooks は、Jupyter Notebook に基づいたマネージド環境です。、Spark SQL、Spark R PySpark、および Scala の Spark マジックカーネルをサポートします。EMR Notebooks は、Amazon EMR リリース 5.18.0 以降を使用して作成されたクラスターで使用できます。詳細については、「Amazon EMR 管理ガイド」の「[EMR Notebooks の使用](#)」を参照してください。
- EMRFS S3 向けに最適化されたコミッターは、Spark と EMRFS を使用して Parquet ファイルを書き込むときに利用できます。このコミッターにより、書き込みパフォーマンスが向上します。詳細については、「[EMRFS S3 向けに最適化されたコミッターを使用する](#)」を参照してください。

## 変更、機能強化、解決した問題

- YARN
  - アプリケーションマスタープロセスをコアノードで実行するように制限するロジックを修正しました。現在、この機能では、yarn-site 設定分類および capacity-scheduler 設定分類の YARN ノードラベル機能とプロパティを使用しています。詳細については、「<https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN>」を参照してください。
- Amazon EMR のデフォルトの Amazon Linux AMI
  - ruby18、php56、および gcc48 は、デフォルトではインストールされなくなりました。必要に応じて、yum を使用してインストールできます。
  - aws-sdk ruby gem は、デフォルトではインストールされなくなりました。必要に応じて、gem install aws-sdk を使用してインストールできます。特定のコンポーネントをインストールすることもできます。例えば gem install aws-sdk-s3 です。

## 既知の問題

- EMR Notebooks — 状況によっては、複数のノートブックエディタが開いているときに、ノートブックエディタがクラスターに接続できないように見えることがあります。このような場合は、ブラウザの Cookie をクリアしてから、ノートブックエディタを再度開きます。
- CloudWatch ContainerPending メトリクスと自動スケーリング — (5.20.0 で修正) Amazon EMR は に負の値を出力することがありますContainerPending。ContainerPending が自動スケーリングルールで使用されている場合、自動スケーリングは期待どおりに動作しません。ContainerPending を自動スケーリングで使用しないでください。



- Amazon EMR バージョン 5.19.0、5.20.0、および 5.21.0 では、YARN ノードラベルは HDFS ディレクトリに保存されます。状況によっては、このために、コアノードの起動が遅延し、クラスターのタイムアウトや起動エラーが発生します。Amazon EMR 5.22.0 以降、この問題は解決されています。YARN ノードラベルは各クラスターノードのローカルディスクに保存され、HDFS への依存が回避されています。

## 5.19.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.2.0	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.5.1	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。

コンポーネント	Version	説明
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.1.0	EMR S3Select コネクタ
emrfs	2.28.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.6.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.5-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.5-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.8.5-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.5-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.5-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.5-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.5-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.5-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.5-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.5-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.7	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.7	1 つ以上の HBase リージョンに対応するサービス。

コンポーネント	Version	説明
hbase-client	1.4.7	HBase コマンドラインクライアント。
hbase-rest-server	1.4.7	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.7	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.3-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.9.4	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.3.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.212	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.212	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.2	Spark コマンドラインクライアント。
spark-history-server	2.3.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	1.11.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.13	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.13	ZooKeeper コマンドラインクライアント。

### 5.19.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.19.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。



分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。

分類	説明
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。

分類	説明
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-memory	Presto の memory.properties ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
presto-connector-tpcds	Presto の tpcds.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.18.1

### 5.18.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
AWS SDK for Java	1.11.393	1.11.393	1.11.336	1.11.336
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.6.0	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.210	0.206	0.206
Spark	2.3.2	2.3.2	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12



## 5.18.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 5.18.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.1.3	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.6.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.5.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.1.0	EMR S3Select コネクタ
emrfs	2.27.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.6.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.4-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。

コンポーネント	Version	説明
hadoop-hdfs-datanode	2.8.4-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.4-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.4-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.4-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.4-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.4-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.4-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.4-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.7	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。

コンポーネント	Version	説明
hbase-region-server	1.4.7	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.7	HBase コマンドラインクライアント。
hbase-rest-server	1.4.7	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.7	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.3-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。

コンポーネント	Version	説明
hive-server2	2.3.3-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.2.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.210	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.210	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.2	Spark コマンドラインクライアント。
spark-history-server	2.3.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.2	YARN スレーブで必要な Apache Spark ライブラリ。

コンポーネント	Version	説明
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.9.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。

### 5.18.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.18.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。



分類	説明
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。

分類	説明
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。

分類	説明
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。

分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。

分類	説明
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.18.0

### 5.18.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
AWS SDK for Java	1.11.393	1.11.336	1.11.336	1.11.336
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.5.2	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.6	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.206	0.206	0.206
Spark	2.3.2	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

## 5.18.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.18.0 に関する情報が含まれています。5.17.0 からの変更が含まれています。



初回リリース日: 2018 年 10 月 24 日

## アップグレード

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2
- Zeppelin 0.8.0

## 新機能

- Amazon EMR 5.18.0 以降では、Amazon EMR アーティファクトリポジトリを使用し、特定の Amazon EMR リリースで利用可能なライブラリおよび依存関係の正確なバージョンに対してジョブコードをビルドできます。詳細については、「[Amazon EMR アーティファクトリポジトリを使用して依存関係を確認する](#)」を参照してください。

## 変更、機能強化、解決した問題

- [Hive]
  - S3 Select のサポートが追加されました。詳細については、「[S3 Select と Hive を使用してパフォーマンスを向上させる](#)」を参照してください。
- Presto
  - [S3 Select Pushdown](#) のサポートが追加されました。詳細については、「[S3 Select Pushdown と Presto を使用してパフォーマンスを向上させる](#)」を参照してください。
- Spark
  - Spark のデフォルトの log4j 設定が、Spark ストリーミングジョブのコンテナログを毎時ロールするように変更されました。これにより、長時間実行されている Spark ストリーミングジョブのログが削除されるのを防ぐことができます。

## 5.18.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通

常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.1.3	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.6.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.5.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emr-s3-select</code>	1.1.0	EMR S3Select コネクタ
<code>emrfs</code>	2.27.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.6.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。

コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.4-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.4-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.4-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.4-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.4-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.4-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.8.4-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.4-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.4-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.7	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.7	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.7	HBase コマンドラインクライアント。
hbase-rest-server	1.4.7	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.7	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.3-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.3-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
nginx	1.12.1	nginx [engine x] は HTTP およびリバースプロキシサーバー

コンポーネント	Version	説明
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.2.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.210	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.210	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.2	Spark コマンドラインクライアント。
spark-history-server	2.3.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.9.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.8.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。

## 5.18.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.18.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。



分類	説明
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。

分類	説明
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。

分類	説明
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。

分類	説明
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.17.2

### 5.17.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.336
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.2	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.6	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0



	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.206	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

## 5.17.2 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 5.17.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.1.3	Amazon SageMaker Spark SDK
emr-ddb	4.6.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.5.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.0.0	EMR S3Select コネクタ
emrfs	2.26.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.5.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれ

コンポーネント	Version	説明
		た、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.4-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.4-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.4-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.4-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.4-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.4-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.8.4-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.4-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.4-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.6	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.6	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.6	HBase コマンドラインクライアント。
hbase-rest-server	1.4.6	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.6	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.3-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.3-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。

コンポーネント	Version	説明
mxnet	1.2.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.206	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.206	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。

コンポーネント	Version	説明
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.1	Spark コマンドラインクライアント。
spark-history-server	2.3.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.9.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。

## 5.17.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.17.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。



分類	説明
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。

分類	説明
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。

分類	説明
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。

分類	説明
phoenix-hbase-metrics	Phoenix の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
phoenix-hbase-site	Phoenix の <code>hbase-site.xml</code> ファイルの値を変更します。
phoenix-log4j	Phoenix の <code>log4j.properties</code> ファイルの値を変更します。
phoenix-metrics	Phoenix の <code>hadoop-metrics2-phoenix.properties</code> ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の <code>pig.properties</code> ファイルの値を変更します。
pig-log4j	Pig の <code>log4j.properties</code> ファイルの値を変更します。
presto-log	Presto の <code>log.properties</code> ファイルの値を変更します。
presto-config	Presto の <code>config.properties</code> ファイルの値を変更します。
presto-password-authenticator	Presto の <code>password-authenticator.properties</code> ファイルの値を変更します。
presto-env	Presto の <code>presto-env.sh</code> ファイルの値を変更します。
presto-node	Presto の <code>node.properties</code> ファイルの値を変更します。
presto-connector-blackhole	Presto の <code>blackhole.properties</code> ファイルの値を変更します。

分類	説明
presto-connector-cassandra	Presto の <code>cassandra.properties</code> ファイルの値を変更します。
presto-connector-hive	Presto の <code>hive.properties</code> ファイルの値を変更します。
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。
presto-connector-kafka	Presto の <code>kafka.properties</code> ファイルの値を変更します。
presto-connector-localfile	Presto の <code>localfile.properties</code> ファイルの値を変更します。
presto-connector-mongodb	Presto の <code>mongodb.properties</code> ファイルの値を変更します。
presto-connector-mysql	Presto の <code>mysql.properties</code> ファイルの値を変更します。
presto-connector-postgresql	Presto の <code>postgresql.properties</code> ファイルの値を変更します。
presto-connector-raptor	Presto の <code>raptor.properties</code> ファイルの値を変更します。
presto-connector-redis	Presto の <code>redis.properties</code> ファイルの値を変更します。
presto-connector-redshift	Presto の <code>redshift.properties</code> ファイルの値を変更します。
presto-connector-tpch	Presto の <code>tpcj.properties</code> ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。

分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.17.1

### 5.17.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.336
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.0	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.4	1.4.4

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.203	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4



	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

### 5.17.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.17.1 に関する情報が含まれています。5.17.0 からの変更が含まれています。

初回リリース日: 2019 年 7 月 18 日

変更点、機能強化、解決した問題

- TCP SACK サービス拒否の問題 ([AWS-2019-005](#)) など、重要な Linux カーネルセキュリティ更新を含めるように、Amazon EMR のデフォルトの Amazon Linux AMI を更新しました。

### 5.17.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.1.3	Amazon SageMaker Spark SDK
emr-ddb	4.6.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.5.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.0.0	EMR S3Select コネクタ
emrfs	2.26.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.5.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.4-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.4-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.4-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.4-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.4-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.4-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.4-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.4-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.4-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.6	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.6	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.6	HBase コマンドラインクライアント。
hbase-rest-server	1.4.6	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.6	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-1	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-hbase	2.3.3-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.2.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.206	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.206	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.1	Spark コマンドラインクライアント。
spark-history-server	2.3.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.3.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tensorflow	1.9.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。

### 5.17.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.17.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。



分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。

分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。

分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の <code>hadoop-metrics2-phoenix.properties</code> ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の <code>pig.properties</code> ファイルの値を変更します。
pig-log4j	Pig の <code>log4j.properties</code> ファイルの値を変更します。
presto-log	Presto の <code>log.properties</code> ファイルの値を変更します。
presto-config	Presto の <code>config.properties</code> ファイルの値を変更します。
presto-password-authenticator	Presto の <code>password-authenticator.properties</code> ファイルの値を変更します。
presto-env	Presto の <code>presto-env.sh</code> ファイルの値を変更します。
presto-node	Presto の <code>node.properties</code> ファイルの値を変更します。
presto-connector-blackhole	Presto の <code>blackhole.properties</code> ファイルの値を変更します。
presto-connector-cassandra	Presto の <code>cassandra.properties</code> ファイルの値を変更します。
presto-connector-hive	Presto の <code>hive.properties</code> ファイルの値を変更します。
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。

分類	説明
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.17.0

### 5.17.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.333
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.0	1.5.0	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.4	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.4.0
MXNet	1.2.0	1.2.0	1.2.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.203	0.203	0.194
Spark	2.3.1	2.3.1	2.3.1	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12



## 5.17.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.17.0 に関する情報が含まれています。5.16.0 からの変更が含まれています。

初回リリース日: 2018 年 8 月 30 日

### アップグレード

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

### 新機能

- TensorFlow のサポートが追加されました。詳細については、「[TensorFlow](#)」を参照してください。

### 変更点、機能強化、解決した問題

- JupyterHub
  - Amazon S3 でのノートブックの永続性のサポートが追加されました。詳細については、「[Amazon S3 でノートブックの永続性を設定するには](#)」を参照してください。
- Spark
  - [S3 Select](#) のサポートが追加されました。詳細については、「[S3 Select と Spark を使用してクエリパフォーマンスを向上させる](#)」を参照してください。
- Amazon EMR バージョン 5.14.0、5.15.0、または 5.16.0 での Cloudwatch メトリクスと自動スケール機能に関する問題が解決されました。

### 既知の問題

- Livy がインストールされた Kerberized クラスターを作成すると、Livy が失敗して、簡易認証が有効になっていないというエラーが表示されます。Livy サーバーを再起動すると、問題が解決されます。回避策として、クラスターの作成時に、プライマリノードで `sudo restart livy-server` を実行するステップを追加します。
- 作成日が 2018-08-11 の Amazon Linux AMI に基づくカスタム Amazon Linux AMI を使用すると、Oozie サーバーの起動に失敗します。Oozie を使用する場合は、作成日が異なる

Amazon Linux AMI ID に基づいてカスタム AMI を作成します。次の AWS CLI コマンドを使用して、2018.03 バージョンのすべての HVM Amazon Linux AMIs のイメージ IDs とリリース日のリストを返すことができます。これにより、ベースとして適切な Amazon Linux AMI を選択できます。を us-west-2 などのリージョン識別子 MyRegion に置き換えます。

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`][[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

## 5.17.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.1.3	Amazon SageMaker Spark SDK
emr-ddb	4.6.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.5.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emr-s3-select	1.0.0	EMR S3Select コネクタ
emrfs	2.26.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.5.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.4-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。

コンポーネント	Version	説明
hadoop-hdfs-datanode	2.8.4-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.4-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.4-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.4-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.4-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.4-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.4-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resource-manager	2.8.4-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.6	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。

コンポーネント	Version	説明
hbase-region-server	1.4.6	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.6	HBase コマンドラインクライアント。
hbase-rest-server	1.4.6	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.6	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.3-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。

コンポーネント	Version	説明
hive-server2	2.3.3-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.2.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。

コンポーネント	Version	説明
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.206	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.206	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.1	Spark コマンドラインクライアント。
spark-history-server	2.3.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tensorflow	1.9.0	TensorFlow 高性能な数値計算のための オープンソースソフトウェアライブラリ。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。

### 5.17.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.17.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。



分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-s3-conf	Jupyter Notebook の S3 の永続性を設定します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。

分類	説明
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。

分類	説明
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。

分類	説明
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.16.1

### 5.16.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
AWS SDK for Java	1.11.336	1.11.336	1.11.333	1.11.333
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.5.0	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-



	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.4.0	0.4.0
MXNet	1.2.0	1.2.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.203	0.194	0.194
Spark	2.3.1	2.3.1	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

### 5.16.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

**⚠ Important**

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.16.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.1.0	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.6.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。

コンポーネント	Version	説明
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.25.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.5.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.4-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.4-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.4-amzn-0	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.8.4-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httfs-server	2.8.4-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.4-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.4-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.4-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.4-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.4-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.4	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.4	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.4	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.4.4	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.4	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.3-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.2.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.203	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.203	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.1	Spark コマンドラインクライアント。
spark-history-server	2.3.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。

### 5.16.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.16.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。



分類	説明
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。

分類	説明
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。

分類	説明
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。

分類	説明
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。

分類	説明
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.16.0

### 5.16.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
AWS SDK for Java	1.11.336	1.11.333	1.11.333	1.11.297
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.2
HCatalog	2.3.3	2.3.3	2.3.3	2.3.2
Hadoop	2.8.4	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.4.0	0.4.0	0.4.0
MXNet	1.2.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	4.3.0
Phoenix	4.14.0	4.13.0	4.13.0	4.13.0



	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.194	0.194	0.194
Spark	2.3.1	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.10

## 5.16.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.16.0 に関する情報が含まれています。5.15.0 からの変更が含まれています。

初回リリース日: 2018 年 7 月 19 日

### アップグレード

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK for Java 1.11.336

- CUDA 9.2
- Redshift JDBC ドライバー 1.2.15.1025

## 変更、機能強化、解決した問題

- HBase
  - [HBASE-20723](#) をバックポートしました。
- Presto
  - LDAP 認証をサポートするための設定変更。詳細については、「[Amazon EMR の Presto に LDAP 認証を使用する](#)」を参照してください。
- Spark
  - Apache Spark バージョン 2.3.1 は Amazon EMR リリース 5.16.0 以降で利用でき、[CVE-2018-8024](#) および [CVE-2018-1334](#) に対処します。Spark の以前のバージョンを Spark バージョン 2.3.1 以降に移行することをお勧めします。

## 既知の問題

- このリリースバージョンでは、c1.medium インスタンスタイプまたは m1.small インスタンスタイプはサポートされていません。これらのインスタンスタイプのいずれかを使用するクラスターは起動しません。回避策として、別のインスタンスタイプを指定するか、別のリリースバージョンを使用します。
- Livy がインストールされた Kerberized クラスターを作成すると、Livy が失敗して、簡易認証が有効になっていないというエラーが表示されます。Livy サーバーを再起動すると、問題が解決されます。回避策として、クラスターの作成時に、プライマリノードで `sudo restart livy-server` を実行するステップを追加します。
- プライマリノードの再起動またはインスタンスコントローラーの再起動後、CloudWatch メトリクスは収集されず、自動スケーリング機能は Amazon EMR バージョン 5.14.0、5.15.0、または 5.16.0 では使用できません。この問題は、Amazon EMR 5.17.0 で修正されています。

## 5.16.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.1.0	Amazon SageMaker Spark SDK
emr-ddb	4.6.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.25.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.5.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム

コンポーネント	Version	説明
		アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.4-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.4-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.4-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.4-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.4-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.4-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.4-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.8.4-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.4-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.4-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.4	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.4	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.4	HBase コマンドラインクライアント。
hbase-rest-server	1.4.4	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.4	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.3-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.5.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.2.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。

コンポーネント	Version	説明
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.2.88	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.14.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.14.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.203	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.203	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト

コンポーネント	Version	説明
spark-client	2.3.1	Spark コマンドラインクライアント。
spark-history-server	2.3.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。



## 5.16.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.16.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。

分類	説明
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。

分類	説明
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。

分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。

分類	説明
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-password-authenticator	Presto の password-authenticator.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。

分類	説明
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。

分類	説明
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.15.1

### 5.15.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
AWS SDK for Java	1.11.333	1.11.333	1.11.297	1.11.297
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.2	1.4.2



	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
HCatalog	2.3.3	2.3.3	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.10	3.4.10

### 5.15.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.15.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.24.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコネクタ。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-https-server	2.8.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resource-manager	2.8.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.4	リージョンの調整および管理コマンドの実行を行う HBase クラスタ用のサービス。
hbase-region-server	1.4.4	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.4	HBase コマンドラインクライアント。
hbase-rest-server	1.4.4	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.4	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.3-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-hbase	2.3.3-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.1.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット

コンポーネント	Version	説明
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.194	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.194	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.0	Spark コマンドラインクライアント。
spark-history-server	2.3.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.3.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。

### 5.15.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。



## emr-5.15.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。

分類	説明
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。

分類	説明
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。

分類	説明
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。

分類	説明
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.15.0

### 5.15.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
AWS SDK for Java	1.11.333	1.11.297	1.11.297	1.11.297
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.2	1.4.2	1.4.2
HCatalog	2.3.3	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-



	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.10	3.4.10	3.4.10

## 5.15.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.15.0 に関する情報が含まれています。5.14.0 からの変更が含まれています。

初回リリース日: 2018 年 6 月 21 日

## アップグレード

- HBase を 1.4.4 にアップグレードしました
- Hive を 2.3.3 にアップグレードしました
- Hue を 4.2.0 にアップグレードしました
- Oozie を 5.0.0 にアップグレードしました
- ZooKeeper を 3.4.12 にアップグレードしました
- AWS SDK を 1.11.333 にアップグレードしました

## 変更点、機能強化、解決した問題

- [Hive]
  - [HIVE-18069](#) をバックポートしました。
- Hue
  - Kerberos が有効になっている場合に Livy で正しく認証されるように Hue を更新しました。Amazon EMR で Kerberos を使用する場合、Livy がサポートされるようになりました。
- JupyterHub
  - Amazon EMR がデフォルトで LDAP クライアントライブラリをインストールする JupyterHub ように更新されました。
  - 自己署名証明書を生成するスクリプトのエラーを修正しました。

## 既知の問題

- このリリースバージョンでは、c1.medium インスタンスタイプまたは m1.small インスタンスタイプはサポートされていません。これらのインスタンスタイプのいずれかを使用するクラスターは起動しません。回避策として、別のインスタンスタイプを指定するか、別のリリースバージョンを使用します。
- プライマリノードの再起動またはインスタンスコントローラーの再起動後、CloudWatch メトリクスは収集されず、自動スケーリング機能は Amazon EMR バージョン 5.14.0、5.15.0、または 5.16.0 では使用できません。この問題は、Amazon EMR 5.17.0 で修正されています。

## 5.15.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.0.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.24.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.4.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	2.8.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.4	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.4	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.4	HBase コマンドラインクライアント。
hbase-rest-server	1.4.4	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.4	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	2.3.3-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.3-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.3-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.3-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.3-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.3-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.3-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.2.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー

コンポーネント	Version	説明
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.1.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	5.0.0	Oozie コマンドラインクライアント。
oozie-server	5.0.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.194	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.194	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.0	Spark コマンドラインクライアント。
spark-history-server	2.3.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。



コンポーネント	Version	説明
zookeeper-server	3.4.12	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.12	ZooKeeper コマンドラインクライアント。

### 5.15.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.15.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。

分類	説明
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。

分類	説明
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。

分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.14.2

### 5.14.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.297
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-



	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.14.2 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 5.14.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオー

プンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.23.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.2	Apache Flink のコマンドラインクライアントスクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計

コンポーネント	Version	説明
		する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.8.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.2	HBase コマンドラインクライアント。
hbase-rest-server	1.4.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.1.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。

コンポーネント	Version	説明
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.194	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.194	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト

コンポーネント	Version	説明
spark-client	2.3.0	Spark コマンドラインクライアント。
spark-history-server	2.3.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。



## 5.14.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.14.2 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。

分類	説明
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。

分類	説明
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。

分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。

分類	説明
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。

分類	説明
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.14.1

### 5.14.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.297
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2



	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.14.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.14.1 に関する情報が含まれています。5.14.0 からの変更が含まれています。

初回リリース日: 2018 年 10 月 17 日

潜在的なセキュリティ脆弱性に対処するために Amazon EMR のデフォルトの AMI を更新しました。

### 5.14.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.0.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。

コンポーネント	Version	説明
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.23.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-1	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.8.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httfs-server	2.8.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.2	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.4.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.1.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.194	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.194	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.0	Spark コマンドラインクライアント。
spark-history-server	2.3.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.14.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.14.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。



分類	説明
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。

分類	説明
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。

分類	説明
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。

分類	説明
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.14.0

### 5.14.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[JupyterHub](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.267
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0



	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
Presto	0.194	0.194	0.194	0.188
Spark	2.3.0	2.3.0	2.3.0	2.2.1
Sqoop	1.4.7	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.14.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.14.0 に関する情報が含まれています。5.13.0 からの変更が含まれています。

初回リリース日: 2018 年 6 月 4 日

### アップグレード

- Apache Flink を 1.4.2 にアップグレードしました
- Apache MXNet を 1.1.0 にアップグレードしました
- Apache Sqoop を 1.4.7 にアップグレードしました

### 新機能

- JupyterHub サポートを追加しました。詳細については、「[JupyterHub](#)」を参照してください。

### 変更、機能強化、解決した問題

- EMRFS

- Amazon S3 へのリクエスト内の UserAgent 文字列が、呼び出し元のプリンシパルのユーザーおよびグループの情報を含めるように更新されました。これは、より包括的なリクエスト追跡のための AWS CloudTrail ログで使用できます。
- HBase
  - [HBASE-20447](#) が含まれました。これにより、特に分割されたリージョンで、キャッシュの問題を引き起こす可能性のある問題に対処します。
- MXNet
  - OpenCV ライブラリを追加しました。
- Spark
  - Spark が EMRFS を使用して Amazon S3 の場所に Parquet ファイルを書き込むと、FileOutputCommitter アルゴリズムがバージョン 1 ではなくバージョン 2 を使用するように更新されました。これにより、名前変更の数が減り、アプリケーションのパフォーマンスが向上します。この変更は、以下には影響しません。
    - Spark 以外のアプリケーション。
    - HDFS などの他のファイルシステムに書き込むアプリケーション (まだのバージョン 1 を使用します FileOutputCommitter ) 。
    - テキストや csv など、EMRFS 直接書き込みを既に使用している、他の出力形式を使用するアプリケーション。

## 既知の問題

- JupyterHub
  - クラスターの作成時に設定分類を使用して Jupyter Notebook JupyterHub と個々の Jupyter Notebook をセットアップすることはサポートされていません。各ユーザーの `jupyterhub_config.py` ファイルと `jupyter_notebook_config.py` ファイルを手動で編集します。詳細については、「[の設定 JupyterHub](#)」を参照してください。
  - JupyterHub はプライベートサブネット内のクラスターで起動に失敗し、メッセージで失敗します `Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'`。これは、自己署名証明書を生成するスクリプトのエラーが原因で発生します。次の回避策を使用して自己署名証明書を生成します。すべてのコマンドは、プライマリノードに接続している間に実行されます。
    1. 証明書生成スクリプトをコンテナからプライマリノードにコピーします。

```
sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
```

2. テキストエディタを使用して 23 行目を変更し、以下に示すようにパブリックホスト名をローカルホスト名に変更します。

```
local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
```

3. スクリプトを実行して、自己署名証明書を生成します。

```
sudo bash ./gen_self_signed_cert.sh
```

4. スクリプトによって生成される証明書ファイルを `/etc/jupyter/conf/` ディレクトリに移動します。:

```
sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
```

`tail jupyter.log` ファイルを使用して、`g` JupyterHub 再起動し、200 レスポンスコードを返していることを確認できます。例:

```
tail -f /var/log/jupyter/jupyter.log
```

これで、次のようなレスポンスが返されます。

```
# [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
https://:9443/
# 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
```

- プライマリノードの再起動またはインスタンスコントローラーの再起動後、CloudWatch メトリクスは収集されず、自動スケーリング機能は Amazon EMR バージョン 5.14.0、5.15.0、または 5.16.0 では使用できません。この問題は、Amazon EMR 5.17.0 で修正されています。

## 5.14.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.23.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.8.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.2	HBase コマンドラインクライアント。
hbase-rest-server	1.4.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
jupyterhub	0.8.1	Jupyter Notebook のマルチユーザーサーバー
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.1.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。

コンポーネント	Version	説明
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
opencv	3.4.0	オープンソースのコンピュータビジョンライブラリ。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.194	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.194	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト



コンポーネント	Version	説明
spark-client	2.3.0	Spark コマンドラインクライアント。
spark-history-server	2.3.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.7	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.14.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.14.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
container-log4j	Hadoop YARN の container-log4j.properties ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。

分類	説明
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。

分類	説明
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。

分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
jupyter-notebook-conf	Jupyter Notebook の jupyter_notebook_config.py ファイルの値を変更します。
jupyter-hub-conf	JupyterHubsの jupyterhub_config.py ファイルの値を変更します。
jupyter-sparkmagic-conf	Sparkmagic の config.json ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。

分類	説明
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。

分類	説明
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.13.1

### 5.13.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)



以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
AWS SDK for Java	1.11.297	1.11.297	1.11.267	1.11.267
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.188	0.188
Spark	2.3.0	2.3.0	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

### 5.13.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.13.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.0.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.22.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.8.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.2	1 つ以上の HBase リージョンに対応するサービス。

コンポーネント	Version	説明
hbase-client	1.4.2	HBase コマンドラインクライアント。
hbase-rest-server	1.4.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.0.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー

コンポーネント	Version	説明
presto-coordinator	0.194	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.194	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.0	Spark コマンドラインクライアント。
spark-history-server	2.3.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.3.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。



コンポーネント	Version	説明
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.13.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.13.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。

分類	説明
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。

分類	説明
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。

分類	説明
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.13.0

### 5.13.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
AWS SDK for Java	1.11.297	1.11.267	1.11.267	1.11.267
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.0	1.4.0	1.4.0



	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.188	0.188	0.188
Spark	2.3.0	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.13.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.13.0 に関する情報が含まれています。5.12.0 からの変更が含まれています。

### アップグレード

- Spark を 2.3.0 にアップグレードしました
- HBase を 1.4.2 にアップグレードしました
- Presto を 0.194 にアップグレードしました
- AWS SDK for Java を 1.11.297 にアップグレードしました

### 変更点、機能強化、解決した問題

- [Hive]
  - [HIVE-15436](#) をバックポートしました。ビューのみを返すように Hive API が拡張されました。

### 既知の問題

- MXNet には現在 OpenCV ライブラリがありません。

## 5.13.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.10.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.22.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム

コンポーネント	Version	説明
		アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.8.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.2	HBase コマンドラインクライアント。
hbase-rest-server	1.4.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-2	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-2	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-2	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-2	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.0.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.194	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.194	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
r	3.4.1	統計的コンピューティング用 R プロジェクト
spark-client	2.3.0	Spark コマンドラインクライアント。
spark-history-server	2.3.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	2.3.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.3.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.13.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。



## emr-5.13.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の <code>log4j2.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-beeline-log4j2	Hive の <code>beeline-log4j2.properties</code> ファイルの値を変更します。
hive-parquet-logging	Hive の <code>parquet-logging.properties</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。

分類	説明
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。

分類	説明
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。

分類	説明
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。

分類	説明
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.12.3

### 5.12.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
AWS SDK for Java	1.11.267	1.11.267	1.11.267	1.11.267

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0



	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.188
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

### 5.12.3 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.12.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通

常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.0.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.9.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.21.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.4.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
<code>ganglia-monitor</code>	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれ

コンポーネント	Version	説明
		た、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.8.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.0	HBase コマンドラインクライアント。
hbase-rest-server	1.4.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.3.2-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.2-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。

コンポーネント	Version	説明
mxnet	1.0.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.188	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.188	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.12.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.12.3 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。



分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の <code>log4j2.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-beeline-log4j2	Hive の <code>beeline-log4j2.properties</code> ファイルの値を変更します。
hive-parquet-logging	Hive の <code>parquet-logging.properties</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。

分類	説明
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。

分類	説明
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。

分類	説明
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。

分類	説明
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.12.2

### 5.12.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
AWS SDK for Java	1.11.267	1.11.267	1.11.267	1.11.238

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
Phoenix	4.13.0	4.13.0	4.13.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.12.2 リリースノート

次のリリースノートには、Amazon EMR リリース 5.12.2 に関する情報が含まれています。5.12.1 からの変更が含まれています。

初回リリース日: 2018 年 8 月 29 日

変更、機能強化、解決した問題

- このリリースでは、潜在的なセキュリティ脆弱性に対処しています。

## 5.12.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ



ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.9.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.21.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム

コンポーネント	Version	説明
		アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.8.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.0	HBase コマンドラインクライアント。
hbase-rest-server	1.4.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.2-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.0.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.188	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.188	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.2.1	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.12.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.12.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。



分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。

分類	説明
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。

分類	説明
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.12.1

### 5.12.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
AWS SDK for Java	1.11.267	1.11.267	1.11.238	1.11.238
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
Presto	0.188	0.188	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.12.1 リリースノート

次のリリースノートには、Amazon EMR リリース 5.12.1 に関する情報が含まれています。5.12.0 からの変更が含まれています。

初回リリース日: 2018 年 3 月 29 日

変更、機能強化、解決した問題

- 潜在的な脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。

## 5.12.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.9.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.21.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。



コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.8.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.8.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.8.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.0	HBase コマンドラインクライアント。
hbase-rest-server	1.4.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.2-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.0.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.188	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.188	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.2.1	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.12.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.12.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。



分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。

分類	説明
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。

分類	説明
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.12.0

### 5.12.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
AWS SDK for Java	1.11.267	1.11.238	1.11.238	1.11.238
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
Presto	0.188	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.12.0 リリースノート

次のリリースノートには、Amazon EMR リリース 5.12.0 に関する情報が含まれています。5.11.1 からの変更が含まれています。

### アップグレード

- AWS SDK for Java 1.11.238 Pub 1.11.267。詳細については、の [AWS SDK for Java 変更ログ](#) を参照してください GitHub。
- Hadoop 2.7.3 ⇒ 2.8.3。詳細については、「[Apache Hadoop リリース](#)」を参照してください。
- Flink 1.3.2 ⇒ 1.4.0。詳細については、「[Apache Flink 1.4.0 リリースに関するお知らせ](#)」を参照してください。
- HBase 1.3.1 ⇒ 1.4.0。詳細については、「[HBase リリースに関するお知らせ](#)」を参照してください。
- Hue 4.0.1 ⇒ 4.1.0。詳細については、「[リリースノート](#)」を参照してください。
- MxNet 0.12.0 ⇒ 1.0.0。詳細については、「[」のMXNet 変更ログ](#)」を参照してください GitHub。
- Presto 0.187 ⇒ 0.188。詳細については、「[リリースノート](#)」を参照してください。

## 変更、機能強化、解決した問題

### • Hadoop

- `yarn.resourcemanager.decommissioning.timeout` プロパティは `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs` に変更されました。このプロパティを使用して、クラスターのスケールダウンをカスタマイズできます。詳細については、「Amazon EMR 管理ガイド」の「[クラスターのスケールダウン](#)」を参照してください。
- Hadoop CLI で、直接コピーを指定する `-d` オプションが `cp (copy)` コマンドに追加されました。これを使用すると、中間の `.COPYING` ファイルを作成しないようにできるため、Amazon S3 間でのデータのコピーが高速になります。詳細については、[HADOOP-12384](#) を参照してください。

### • Pig

- Pig 環境プロパティの設定を簡素化する `pig-env` 構成分類が追加されました。詳細については、「[アプリケーションの設定](#)」を参照してください。

### • Presto

- Presto の `presto-connector-redshift` 設定ファイルで値を指定するために使用できる、`redshift.properties` 設定分類が追加されました。詳細については、Presto のドキュメントの「[Redshift Connector](#)」、および「[アプリケーションの設定](#)」を参照してください。
- EMRFS の Presto サポートが追加されました。これがデフォルト設定です。以前の Amazon EMR リリースでは、唯一のオプションである `PrestoS3FileSystem` が使用されていました。詳細については、「[EMRFS と PrestoS3FileSystem の設定](#)」を参照してください。

#### Note

Amazon EMR バージョン 5.12.0 を使用して Amazon S3 で基盤となるデータをクエリするとき、Presto エラーが発生することがあります。これは、Presto が `emrfs-site.xml` から設定分類値を取得できないためです。回避策として、`usr/lib/presto/plugin/hive-hadoop2/` の下に `emrfs` サブディレクトリを作成し、既存の `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` ファイルへのシンボリックリンクを `usr/lib/presto/plugin/hive-hadoop2/emrfs` に作成します。次に `presto-server` プロセスを再起動します (`sudo presto-server stop` に続いて `sudo presto-server start`)。

### • Spark

- バックポートされた [SPARK-22036: BigDecimal multiplication は、null を返すことがあります](#)。

## 既知の問題

- MXNet には OpenCV ライブラリは含まれていません。
- デフォルトでは R はクラスターノードにインストールされないため、カスタム AMI を使用して作成されたクラスターで SparkR を使用することはできません。

## 5.12.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.0.1	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。



コンポーネント	Version	説明
emr-s3-dist-cp	2.9.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.21.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.4.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.8.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.8.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.8.3-amzn-0	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.8.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httfs-server	2.8.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.8.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.8.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.8.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.8.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.4.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.4.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.4.0	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.4.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.4.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-1	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-1	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.2-amzn-1	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-1	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-1	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.1.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	1.0.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.1.85	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.13.0-HBase-1.4	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.13.0-HBase-1.4	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー

コンポーネント	Version	説明
presto-coordinator	0.188	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.188	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.12.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.12.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。

分類	説明
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。

分類	説明
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。



分類	説明
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-env	Pig 環境の値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。

分類	説明
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-redshift	Presto の redshift.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.11.4

### 5.11.4 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.238
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.11.4 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 5.11.4 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.0	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.8.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.20.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。



コンポーネント	Version	説明
hadoop-hdfs-library	2.7.3-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-6	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-6	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。

コンポーネント	Version	説明
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	0.12.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.0.176	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー

コンポーネント	Version	説明
presto-coordinator	0.187	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.187	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.11.4 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.11.4 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。

分類	説明
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。

分類	説明
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。

分類	説明
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。



分類	説明
phoenix-metrics	Phoenix の <code>hadoop-metrics2-phoenix.properties</code> ファイルの値を変更します。
pig-properties	Pig の <code>pig.properties</code> ファイルの値を変更します。
pig-log4j	Pig の <code>log4j.properties</code> ファイルの値を変更します。
presto-log	Presto の <code>log.properties</code> ファイルの値を変更します。
presto-config	Presto の <code>config.properties</code> ファイルの値を変更します。
presto-env	Presto の <code>presto-env.sh</code> ファイルの値を変更します。
presto-node	Presto の <code>node.properties</code> ファイルの値を変更します。
presto-connector-blackhole	Presto の <code>blackhole.properties</code> ファイルの値を変更します。
presto-connector-cassandra	Presto の <code>cassandra.properties</code> ファイルの値を変更します。
presto-connector-hive	Presto の <code>hive.properties</code> ファイルの値を変更します。
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。
presto-connector-kafka	Presto の <code>kafka.properties</code> ファイルの値を変更します。
presto-connector-localfile	Presto の <code>localfile.properties</code> ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.11.3

### 5.11.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.238
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

### 5.11.3 リリースノート

次のリリースノートには、Amazon EMR リリース 5.11.3 に関する情報が含まれています。5.11.2 からの変更が含まれています。

初回リリース日: 2019 年 7 月 18 日

## 変更点、機能強化、解決した問題

- TCP SACK サービス拒否の問題 ([AWS-2019-005](#)) など、重要な Linux カーネルセキュリティ更新を含めるように、Amazon EMR のデフォルトの Amazon Linux AMI を更新しました。

### 5.11.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.0	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。
<code>emr-s3-dist-cp</code>	2.8.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.20.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-httfs-server	2.7.3-amzn-6	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-6	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。



コンポーネント	Version	説明
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション

コンポーネント	Version	説明
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	0.12.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.0.176	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.187	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.187	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.11.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.11.3 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。

分類	説明
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。

分類	説明
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。

分類	説明
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。

分類	説明
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。



分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。

分類	説明
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.11.2

### 5.11.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.221

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.11.2 リリースノート

次のリリースノートには、Amazon EMR リリース 5.11.2 に関する情報が含まれています。5.11.1 からの変更が含まれています。

初回リリース日: 2018 年 8 月 29 日

変更、機能強化、解決した問題

- このリリースでは、潜在的なセキュリティ脆弱性に対処しています。

## 5.11.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.0	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.8.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.20.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム

コンポーネント	Version	説明
		アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-6	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.7.3-amzn-6	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	0.12.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。



コンポーネント	Version	説明
nvidia-cuda	9.0.176	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.187	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.187	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.2.1	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.11.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.11.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。

分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。

分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。



## Amazon EMR リリース 5.11.1

### 5.11.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
AWS SDK for Java	1.11.238	1.11.238	1.11.221	1.11.221
Python	2.7、3.4	2.7、3.4	2.7、3.4	2.7、3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
HCatalog	2.3.2	2.3.2	2.3.1	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.1	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

### 5.11.1 リリースノート

次のリリースノートには、Amazon EMR 5.11.1 リリースに関する情報が含まれています。Amazon EMR 5.8.0 リリースからの変更が含まれています。

初回リリース日: 2018 年 1 月 22 日

#### 変更、機能強化、解決した問題

- 投機的実行 (CVE-2017-5715、CVE-2017-5753、CVE-2017-5754) に関連する脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。詳細については、「<https://aws.amazon.com/security/security-bulletins/AWS-2018-013/>」を参照してください。

### 5.11.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
aws-sagemaker-spark-sdk	1.0	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.8.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.20.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコネクタ。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-6	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-6	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.2-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-hbase	2.3.2-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	0.12.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.0.176	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。

コンポーネント	Version	説明
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.187	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.187	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。



コンポーネント	Version	説明
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.11.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.11.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。

分類	説明
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。

分類	説明
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。

分類	説明
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。

分類	説明
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。

分類	説明
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.11.0

### 5.11.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
AWS SDK for Java	1.11.238	1.11.221	1.11.221	1.11.183
Python	2.7、3.4	2.7、3.4	2.7、3.4	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.1	2.3.1	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.1	2.3.1	2.3.0



	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.184
Spark	2.2.1	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.11.0 リリースノート

次のリリースノートには、Amazon EMR 5.11.0 に関する情報が含まれています。5.10.0 からの変更が含まれています。

### アップグレード

- Hive 2.3.2
- Spark 2.2.1
- SDK for Java 1.11.238

### 新機能

- Spark
  - `spark.decommissioning.timeout.threshold` 設定が追加されました。スポットインスタンス使用時の Spark 廃棄の動作が向上します。詳細については、「[ノード停止の動作設定](#)」を参照してください。
  - `aws-sagemaker-spark-sdk` コンポーネントを Spark に追加しました。これにより、Amazon Spark と SageMaker Spark の統合に関連する依存関係がインストールされます [SageMaker](#)。Amazon SageMaker Spark を使用して、Amazon SageMaker ステージを使用して Spark 機械学習 (ML) パイプラインを構築できます。詳細については、「Amazon SageMaker デベロッパーガイド」の「での [SageMaker Spark Readme](#) GitHub」と「[Amazon での Apache Spark の使用 SageMaker](#)」を参照してください。

### 既知の問題

- MXNet には OpenCV ライブラリは含まれていません。
- Hive 2.3.2 では `hive.compute.query.using.stats=true` がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、`hive.compute.query.using.stats=true` が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して `SELECT COUNT(*)` クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、`ANALYZE TABLE` コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

## 5.11.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>aws-sagemaker-spark-sdk</code>	1.0	Amazon SageMaker Spark SDK
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.8.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.20.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-6	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-6	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-6	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-6	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-6	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	2.7.3-amzn-6	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-6	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-6	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-6	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	2.3.2-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.2-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.2-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.2-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.2-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.2-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。

コンポーネント	Version	説明
mxnet	0.12.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.0.176	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.187	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.187	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.1	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.2.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.11.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。



## emr-5.11.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の <code>log4j2.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-beeline-log4j2	Hive の <code>beeline-log4j2.properties</code> ファイルの値を変更します。
hive-parquet-logging	Hive の <code>parquet-logging.properties</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。

分類	説明
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。

分類	説明
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。

分類	説明
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

分類	説明
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.10.1

### 5.10.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
AWS SDK for Java	1.11.221	1.11.221	1.11.183	1.11.183
Python	2.7、3.4	2.7、3.4	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.1	2.3.1	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.1	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0



	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
Presto	0.187	0.187	0.184	0.184
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

### 5.10.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.10.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.7.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.20.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計

コンポーネント	Version	説明
		する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-5	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-5	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-5	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-5	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-5	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-5	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-5	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-5	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.7.3-amzn-5	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-5	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	0.12.0	フレキシブルかつスケーラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
nvidia-cuda	9.0.176	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.187	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.187	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.0	Spark コマンドラインクライアント。
spark-history-server	2.2.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.0	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.2.0	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.10.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.10.1 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。



分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。

分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。

分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.10.0

### 5.10.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[MXNet](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
AWS SDK for Java	1.11.221	1.11.183	1.11.183	1.11.160
Python	2.7、3.4	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.1	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	-
MXNet	0.12.0	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.16.0
Presto	0.187	0.184	0.184	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.10.0 リリースノート

次のリリースノートには、Amazon EMR バージョン 5.10.0 リリースに関する情報が含まれています。Amazon EMR 5.9.0 リリースからの変更が含まれています。

### アップグレード

- AWS SDK for Java 1.11.221
- Hive 2.3.1
- Presto 0.187

### 新機能

- Kerberos 認証のサポートが追加されました。詳しくは、「Amazon EMR 管理ガイド」の「[Kerberos 認証を使用する](#)」を参照してください。
- EMRFS の IAM ロールのサポートが追加されました。詳細については、「Amazon EMR 管理ガイド」の「[Amazon S3 への EMRFS リクエストの IAM ロールを設定する](#)」を参照してください。
- GPU ベースの P2 および P3 のインスタンスタイプのサポートが追加されました。詳細については、「[Amazon EC2 P2 インスタンス](#)」および「[Amazon EC2 P3 インスタンス](#)」を参照してください。これらのインスタンスタイプには、デフォルトで NVIDIA ドライバー 384.81 および CUDA ドライバー 9.0.176 がインストールされています。
- [Apache MXNet](#) のサポートが追加されました。

### 変更、機能強化、解決した問題

- Presto



- デフォルトの Hive メタストアとして AWS Glue データカタログを使用するためのサポートが追加されました。詳細については、「[AWS Glue データカタログでの Presto の使用](#)」を参照してください。
- 「[地理空間機能](#)」のサポートが追加されました。
- 結合で「[ディスクに書き出す](#)」ことができるようになりました。
- 「[Redshift コネクタ](#)」のサポートが追加されました。
- Spark
  - 「[SPARK-20640](#)」をバックポートしました。これにより、spark.shuffle.registration.timeout プロパティおよび spark.shuffle.registration.maxAttempts プロパティを使用して、rpc タイムアウトや、シャッフル登録値の再試行回数が設定可能になりました。
  - [SPARK-21549](#) をバックポートしました。これにより、HDFS 以外の場所にカスタム OutputFormat を書き込むときに発生するエラーが修正されます。
  - 「[Hadoop-13270](#)」をバックポートしました。
- Numpy、Scipy、Matplotlib の各ライブラリは、Amazon EMR の基本 AMI から削除されています。アプリケーションでこれらのライブラリが必要な場合、アプリケーションリポジトリで使用できるため、ブートストラップアクションで yum install を使用してすべてのノードにインストールすることができます。
- Amazon EMR の基本 AMI にアプリケーションの RPM パッケージが含まれなくなったため、その RPM パッケージはクラスターノードに存在しなくなりました。カスタム AMI と Amazon EMR の基本 AMI で、Amazon S3 の RPM パッケージリポジトリを参照できるようになりました。
- Amazon EC2 で 1 秒単位の請求が導入されたため、デフォルトの [スケールダウン動作] は [インスタンス時間で削除する] ではなく [タスク完了で削除する] になっています。詳細については、「[クラスターのスケールダウンを設定する](#)」を参照してください。

## 既知の問題

- MXNet には OpenCV ライブラリは含まれていません。
- Hive 2.3.1 では hive.compute.query.using.stats=true がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、hive.compute.query.using.stats=true が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して SELECT COUNT(\*) クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、ANALYZE TABLE コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

## 5.10.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.5.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.7.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.20.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-5	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-5	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-5	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-5	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-httfs-server	2.7.3-amzn-5	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-5	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-5	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-5	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-5	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-5	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。

コンポーネント	Version	説明
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション

コンポーネント	Version	説明
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mxnet	0.12.0	フレキシブルかつスケラブルで、ディープラーニングにおいて効率的なライブラリです。
mysql-server	5.5.54+	MySQL データベースサーバー。
nvidia-cuda	9.0.176	Nvidia ドライバーと Cuda ツールキット
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.187	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.187	いくつかのクエリを実行するサービス。

コンポーネント	Version	説明
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.0	Spark コマンドラインクライアント。
spark-history-server	2.2.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.3	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.10.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.10.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。



分類	説明
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。

分類	説明
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。

分類	説明
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。

分類	説明
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。

分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。

分類	説明
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.9.1

### 5.9.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
AWS SDK for Java	1.11.183	1.11.183	1.11.160	1.11.160

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.16.0	0.16.0
Presto	0.184	0.184	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.9.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 5.9.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通



常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.4.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.7.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.19.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
<code>ganglia-monitor</code>	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-4	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.7.3-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.0-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.0-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.184	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.184	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.0	Spark コマンドラインクライアント。
spark-history-server	2.2.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.9.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.9.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。

分類	説明
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。

分類	説明
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。



分類	説明
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.9.0

### 5.9.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Livy](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
AWS SDK for Java	1.11.183	1.11.160	1.11.160	1.11.160
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	3.12.0	3.12.0	3.12.0

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.16.0	0.16.0	0.16.0
Presto	0.184	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.9.0 リリースノート

次のリリースノートには、Amazon EMR バージョン 5.9.0 リリースに関する情報が含まれています。Amazon EMR 5.8.0 リリースからの変更が含まれています。

リリース日: 2017 年 10 月 5 日

最新機能更新日: 2017 年 10 月 12 日

### アップグレード

- AWS SDK for Java バージョン 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

### 新機能

- Livy サポート (バージョン 0.4.0-incubating) を追加しました。詳細については、「[Apache Livy](#)」を参照してください。
- Spark の Hue ノートブックのサポートを追加しました。
- i3-シリーズ Amazon EC2 インスタンスのサポートを追加しました (2017 年 10 月 12 日)。

### 変更、機能強化、解決した問題

- Spark
  - 手動のサイズ変更または自動のスケーリングポリシーのリクエストによるノードの終了処理を、Spark がより適切に行う、新しい機能のセットを追加しました。詳細については、「[ノード停止の動作設定](#)」を参照してください。
  - ブロック転送サービスの転送時の暗号化には 3DES に代わり SSL を使用します。これにより AES-NI で Amazon EC2 インスタンスタイプを使用する場合にパフォーマンスが向上します。
  - [SPARK-21494](#) を移植しました。
- Zeppelin
  - [ZEPPELIN-2377](#) を移植しました。
- HBase

- パッチ [HBASE-18533](#) を追加しました。これにより、設定分類を使用して HBase BucketCache `hbase-site`設定に追加の値を指定できます。
- Hue
  - Hue の Hive クエリエディタに対する AWS Glue Data Catalog サポートが追加されました。
  - デフォルトでは、Hue のスーパーユーザーは、Amazon EMR の IAM ロールがアクセスを許可されているすべてのファイルにアクセスできます。新しく作成されたユーザーには、Amazon S3 ファイルブラウザへのアクセス許可は自動的に付与されません。グループに対して `filebrowser.s3_access` アクセス許可を有効にする必要があります。
- AWS Glue Data Catalog を使用して作成された基盤となる JSON データにアクセスできない問題を解決しました。

## 既知の問題

- すべてのアプリケーションがインストールされ、デフォルトの Amazon EBS ルートボリュームサイズが変更されていない場合、クラスターの起動は失敗します。回避策として、`aws emr create-cluster` コマンドを使用し AWS CLI、より大きな `--ebs-root-volume-size` パラメータを指定します。
- Hive 2.3.0 では `hive.compute.query.using.stats=true` がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、`hive.compute.query.using.stats=true` が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して `SELECT COUNT(*)` クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、`ANALYZE TABLE` コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

## 5.9.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ



ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.4.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.7.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.19.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.2	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-4	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.7.3-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.0-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.0-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	4.0.1	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
livy-server	0.4.0-incubating	Apache Spark を操作するための REST インターフェイス
mahout-client	0.13.0	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.184	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.184	いくつかのクエリを実行するサービス。
pig-client	0.17.0	Pig コマンドラインクライアント。
spark-client	2.2.0	Spark コマンドラインクライアント。
spark-history-server	2.2.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.9.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.9.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。

分類	説明
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。

分類	説明
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。



分類	説明
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
livy-conf	Livy の livy.conf ファイルの値を変更します。
livy-env	Livy 環境の値を変更します。
livy-log4j	Livy の log4j.properties の設定を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.8.3

### 5.8.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
AWS SDK for Java	1.11.160	1.11.160	1.11.160	1.11.160
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

### 5.8.3 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.8.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.4.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。

コンポーネント	Version	説明
emr-s3-dist-cp	2.6.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.18.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-3	HDFS コマンドラインクライアントとライブラリ



コンポーネント	Version	説明
hadoop-hdfs-namenode	2.7.3-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-ftpfs-server	2.7.3-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.0-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.0-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.13.0	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-1	Pig コマンドラインクライアント。
spark-client	2.2.0	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.2.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.8.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.8.3 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の <code>log4j2.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-beeline-log4j2	Hive の <code>beeline-log4j2.properties</code> ファイルの値を変更します。
hive-parquet-logging	Hive の <code>parquet-logging.properties</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。



分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。

分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.8.2

### 5.8.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
AWS SDK for Java	1.11.160	1.11.160	1.11.160	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.8.2 リリースノート

次のリリースノートには、Amazon EMR リリース 5.8.2 に関する情報が含まれています。5.8.1 からの変更が含まれています。

初回リリース日: 2018 年 3 月 29 日

### 変更、機能強化、解決した問題

- 潜在的な脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。

## 5.8.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	4.4.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.6.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.18.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリ

コンポーネント	Version	説明
		クスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.0-amzn-0	Hive コマンドラインクライアント。



コンポーネント	Version	説明
hive-hbase	2.3.0-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.13.0	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー

コンポーネント	Version	説明
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-1	Pig コマンドラインクライアント。
spark-client	2.2.0	Spark コマンドラインクライアント。
spark-history-server	2.2.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.8.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.8.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。

分類	説明
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。

分類	説明
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。

分類	説明
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。

分類	説明
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。

分類	説明
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。



分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.8.1

### 5.8.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
AWS SDK for Java	1.11.160	1.11.160	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.0	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

### 5.8.1 リリースノート

次のリリースノートには、Amazon EMR 5.8.1 リリースに関する情報が含まれています。Amazon EMR 5.8.0 リリースからの変更が含まれています。

初回リリース日: 2018 年 1 月 22 日

#### 変更、機能強化、解決した問題

- 投機的実行 (CVE-2017-5715、CVE-2017-5753、CVE-2017-5754) に関連する脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しまし

た。詳細については、「<https://aws.amazon.com/security/security-bulletins/AWS-2018-013/>」を参照してください。

### 5.8.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.4.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.6.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.18.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.3.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-3	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	2.7.3-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	2.3.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.0-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.0-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.13.0	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-1	Pig コマンドラインクライアント。
spark-client	2.2.0	Spark コマンドラインクライアント。
spark-history-server	2.2.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.0	YARN スレーブで必要な Apache Spark ライブラリ。



コンポーネント	Version	説明
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.8.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.8.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。

分類	説明
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。

分類	説明
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。

分類	説明
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.8.0

### 5.8.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
AWS SDK for Java	1.11.160	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.0	1.3.0	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.0
HCatalog	2.3.0	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0



	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.1.1	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.8.0 リリースノート

次のリリースノートには、Amazon EMR バージョン 5.8.0 リリースに関する情報が含まれています。Amazon EMR 5.7.0 リリースからの変更が含まれています。

初回リリース日: 2017 年 8 月 10 日

最新機能更新日: 2017 年 9 月 25 日

### アップグレード

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0。詳細については、Apache Hive サイトの「[リリースノート](#)」を参照してください。
- Spark 2.2.0。詳細については、Apache Spark サイトの「[リリースノート](#)」を参照してください。

### 新機能

- アプリケーション履歴の表示のサポートを追加しました (2017 年 9 月 25 日)。詳細については、「Amazon EMR 管理ガイド」の「[アプリケーションの履歴を表示する](#)」を参照してください。

### 変更点、機能強化、解決した問題

- AWS Glue Data Catalog との統合
  - Hive と Spark SQL が AWS Glue Data Catalog を Hive メタデータストアとして使用する機能を追加しました。詳細については、「[AWS Glue データカタログを Hive のメタストアとして使用する](#)」および「[AWS Glue データカタログを Spark SQL のメタストアとして使用する](#)」を参照してください。
- クラスター詳細の [Application history] を追加しました。これにより、YARN アプリケーションの履歴データや、Spark アプリケーションの追加の詳細を表示できます。詳細については、「Amazon EMR 管理ガイド」の「[アプリケーションの履歴を表示する](#)」を参照してください。
- Oozie
  - [OOZIE-2748](#) をバックポートしました。
- Hue
  - [HUE-5859](#) をバックポートしました。
- HBase

- `getMasterInitializedTime` を使用して Java Management Extensions (JMX) 経由で HBase マスターサーバーの開始時間を公開するパッチを追加しました。
- クラスターの開始時間を改善するパッチを追加しました。

## 既知の問題

- すべてのアプリケーションがインストールされ、デフォルトの Amazon EBS ルートボリュームサイズが変更されていない場合、クラスターの起動は失敗します。回避策として、`aws emr create-cluster` コマンドを使用し AWS CLI、より大きな `--ebs-root-volume-size` パラメータを指定します。
- Hive 2.3.0 では `hive.compute.query.using.stats=true` がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、`hive.compute.query.using.stats=true` が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して `SELECT COUNT(*)` クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、`ANALYZE TABLE` コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

- Spark - Spark を使用する場合、`appusher` デーモンには、長時間実行されている Spark ジョブで数時間または数日後に発生する可能性があるファイルハンドラのリークの問題があります。この問題を修正するには、プライマリノードに接続し、「`sudo /etc/init.d/appusher stop`」と入力します。これにより、その `appusher` デーモンが停止し、Amazon EMR がこのデーモンを自動的に再起動します。
- アプリケーションの履歴
  - Spark のデッドエグゼキュターの履歴データは利用できません。
  - アプリケーション履歴は、セキュリティ設定を使用してインフライト暗号化を有効にするクラスターでは利用できません。

## 5.8.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通

常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.4.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.4.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.4.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.6.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.18.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.3.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
<code>ganglia-monitor</code>	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-3	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.7.3-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.3.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.3.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.3.0-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.3.0-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.3.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.3.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.13.0	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-1	Pig コマンドラインクライアント。
spark-client	2.2.0	Spark コマンドラインクライアント。
spark-history-server	2.2.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.2.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.2.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。



コンポーネント	Version	説明
zeppelin-server	0.7.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.8.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.8.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。

分類	説明
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の <code>hadoop-metrics2-phoenix.properties</code> ファイルの値を変更します。
pig-properties	Pig の <code>pig.properties</code> ファイルの値を変更します。
pig-log4j	Pig の <code>log4j.properties</code> ファイルの値を変更します。
presto-log	Presto の <code>log.properties</code> ファイルの値を変更します。
presto-config	Presto の <code>config.properties</code> ファイルの値を変更します。
presto-env	Presto の <code>presto-env.sh</code> ファイルの値を変更します。
presto-node	Presto の <code>node.properties</code> ファイルの値を変更します。
presto-connector-blackhole	Presto の <code>blackhole.properties</code> ファイルの値を変更します。
presto-connector-cassandra	Presto の <code>cassandra.properties</code> ファイルの値を変更します。
presto-connector-hive	Presto の <code>hive.properties</code> ファイルの値を変更します。
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。
presto-connector-kafka	Presto の <code>kafka.properties</code> ファイルの値を変更します。
presto-connector-localfile	Presto の <code>localfile.properties</code> ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.7.1

### 5.7.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.0	1.3.0	1.2.1	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-



	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.7.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

**⚠ Important**

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.7.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。
<code>emr-s3-dist-cp</code>	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.18.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.3.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-httfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。

コンポーネント	Version	説明
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.13.0	機械学習用のライブラリ。

コンポーネント	Version	説明
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.1	Spark コマンドラインクライアント。
spark-history-server	2.1.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.1	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.1.1	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.7.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.7.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。



分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。

分類	説明
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。

分類	説明
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.7.0

### 5.7.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.0	1.2.1	1.2.1	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
HBase	1.3.1	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.7.0 リリースノート

次のリリースノートには、Amazon EMR 5.7.0 に関する情報が含まれています。Amazon EMR 5.6.0 リリースからの変更が含まれています。

リリース日: 2017 年 7 月 13 日

### アップグレード

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

### 新機能

- クラスターの作成時に、カスタム Amazon Linux AMI を指定する機能を追加しました。詳細については、「[カスタム AMI の使用](#)」を参照してください。

### 変更、機能強化、解決した問題

- HBase
  - HBase のリードレプリカクラスターを設定する機能を追加しました。詳細については、「[リードレプリカクラスターの使用](#)」を参照してください。
  - 複数のバグ修正と機能強化
- Presto – node.properties を設定する機能を追加しました。
- YARN – container-log4j.properties を設定する機能を追加しました。
- Sqoop – [SQOOP-2880](#) をバックポートしました。Sqoop 一時ディレクトリを設定できる引数が導入されています。



## 5.7.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.18.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.3.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。

コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.13.0	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。

コンポーネント	Version	説明
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.11.0-HBase-1.3	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.11.0-HBase-1.3	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.1	Spark コマンドラインクライアント。
spark-history-server	2.1.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.7.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.7.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。

分類	説明
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。



分類	説明
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。

分類	説明
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。

分類	説明
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。

分類	説明
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.6.1

### 5.6.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.1	1.2.1	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.6.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

**⚠ Important**

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.6.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。
<code>emr-s3-dist-cp</code>	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.17.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.2.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。



コンポーネント	Version	説明
hadoop-httfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。

コンポーネント	Version	説明
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.13.0	機械学習用のライブラリ。

コンポーネント	Version	説明
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.1	Spark コマンドラインクライアント。
spark-history-server	2.1.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.1	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.1.1	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.6.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.6.1 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。

分類	説明
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。



分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。

分類	説明
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.6.0

### 5.6.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.1	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.6.0 リリースノート

次のリリースノートには、Amazon EMR 5.6.0 リリースに関する情報が含まれています。Amazon EMR 5.5.0 リリースからの変更が含まれています。

リリース日: 2017 年 6 月 5 日

### アップグレード

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0。Amazon EMR バージョン 5.0 以降の Spark 2.x をサポートする最初のバージョンの Mahout です。
- Spark 2.1.1

### 変更、機能強化、解決した問題

- Presto
  - セキュリティ設定を使って転送時の暗号化を有効にして、Presto ノード間で SSL/TLS を使った安全な通信を有効にする機能が追加されました。詳細については、「[伝送時のデータ暗号化](#)」を参照してください。
  - [Presto 7661](#) をバックポートしました。これにより VERBOSE オプションを EXPLAIN ANALYZE ステートメントに追加し、クエリプランについてのより詳細なレポートと低レベルの統計を作成できます。

## 5.6.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.17.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.2.1	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。

コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.1	HBase コマンドラインクライアント。
hbase-rest-server	1.3.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。



コンポーネント	Version	説明
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.13.0	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。

コンポーネント	Version	説明
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.1	Spark コマンドラインクライアント。
spark-history-server	2.1.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。

コンポーネント	Version	説明
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.6.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.6.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。

分類	説明
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。

分類	説明
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。

分類	説明
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-env	Presto の presto-env.sh ファイルの値を変更します。
presto-node	Presto の node.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。

分類	説明
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。



分類	説明
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.5.4

### 5.5.4 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

## 5.5.4 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

**⚠ Important**

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 5.5.4 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.16.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.2.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-httfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.0	HBase コマンドラインクライアント。
hbase-rest-server	1.3.0	HBase の RESTful HTTP エンドポイントを提供するサービス。

コンポーネント	Version	説明
hbase-thrift-server	1.3.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。

コンポーネント	Version	説明
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。



コンポーネント	Version	説明
spark-yarn-slave	2.1.0	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.5.4 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.5.4 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。

分類	説明
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.5.3

### 5.5.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0



	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

### 5.5.3 リリースノート

次のリリースノートには、Amazon EMR リリース 5.5.3 に関する情報が含まれています。5.5.2 からの変更が含まれています。

初回リリース日: 2018 年 8 月 29 日

変更、機能強化、解決した問題

- このリリースでは、潜在的なセキュリティ脆弱性に対処しています。

### 5.5.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-*amzn*-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。

コンポーネント	Version	説明
emr-s3-dist-cp	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.16.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.2.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.0	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.3.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.5.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.5.3 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。



分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の <code>log4j2.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-beeline-log4j2	Hive の <code>beeline-log4j2.properties</code> ファイルの値を変更します。
hive-parquet-logging	Hive の <code>parquet-logging.properties</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。

分類	説明
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.5.2

### 5.5.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.9

## 5.5.2 リリースノート

次のリリースノートには、Amazon EMR リリース 5.5.2 に関する情報が含まれています。5.5.1 からの変更が含まれています。

初回リリース日: 2018 年 3 月 29 日

### 変更、機能強化、解決した問題

- 潜在的な脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。

## 5.5.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。



コンポーネント	Version	説明
emr-ddb	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.16.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.2.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリ

コンポーネント	Version	説明
		クスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.0	HBase コマンドラインクライアント。
hbase-rest-server	1.3.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー

コンポーネント	Version	説明
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

コンポーネント	Version	説明
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.5.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.5.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。

分類	説明
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。

分類	説明
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。



分類	説明
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。

分類	説明
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。

分類	説明
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。

分類	説明
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.5.1

### 5.5.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Span](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.166	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.0	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.9	3.4.9

### 5.5.1 リリースノート

次のリリースノートには、Amazon EMR 5.5.1 リリースに関する情報が含まれています。Amazon EMR 5.5.0 リリースからの変更が含まれています。

初回リリース日: 2018 年 1 月 22 日

変更、機能強化、解決した問題

- 投機的実行 (CVE-2017-5715、CVE-2017-5753、CVE-2017-5754) に関連する脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。詳細については、「<https://aws.amazon.com/security/security-bulletins/AWS-2018-013/>」を参照してください。

### 5.5.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通

常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.16.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.2.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
<code>ganglia-monitor</code>	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。



コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.0	HBase コマンドラインクライアント。
hbase-rest-server	1.3.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.7.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

### 5.5.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.5.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。

分類	説明
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。



分類	説明
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。

分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.5.0

### 5.5.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.166	0.166	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.0	0.7.0	0.6.2
ZooKeeper	3.4.10	3.4.9	3.4.9	3.4.9

## 5.5.0 リリースノート

次のリリースノートには、Amazon EMR 5.5.0 リリースに関する情報が含まれています。Amazon EMR 5.4.0 リリースからの変更が含まれています。

リリース日: 2017 年 4 月 26 日

### アップグレード

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1

- ZooKeeper 3.4.10

## 変更点、機能強化、解決した問題

- Spark

- Spark のパッチ「[\(SPARK-20115\) 外部シャッフルサービスが利用できない場合に、すべての失われたシャッフルブロックを再計算するように DAGScheduler を修正](#)」を Spark バージョン 2.1.0 にしました。これはこのリリースに含まれています。

- Flink

- Flink は Scala 2.11 で作成されるようになりました。プロジェクトで Scala API とライブラリを使用する場合は、Scala 2.11 を使用することをお勧めします。
- HADOOP\_CONF\_DIR と YARN\_CONF\_DIR のデフォルトが適切に設定されないため start-scala-shell.sh が機能しない問題に対応しました。さらに env.hadoop.conf.dir または env.yarn.conf.dir 設定分類の /etc/flink/conf/flink-conf.yaml と flink-conf を使って、これらの値を設定する機能を追加しました。
- EMR 固有の新しいコマンドで、flink-scala-shell のラッパーとなる start-scala-shell.sh を追加しました。start-scala-shell に代えて、このコマンドを使用することをお勧めします。新しいコマンドにより実行が簡素化されます。たとえば、flink-scala-shell -n 2 は、タスクの並行度 2 で、Flink Scala シェルを開始します。
- EMR 固有の新しいコマンドで、flink-yarn-session のラッパーとなる yarn-session.sh を追加しました。yarn-session に代えて、このコマンドを使用することをお勧めします。新しいコマンドにより実行が簡素化されます。たとえば、flink-yarn-session -d -n 2 は長時間稼働の Flink セッションを、デタッチ状態で、2 つのタスクマネージャを使って開始します。
- 「[\(FLINK-6125\) Commons httpClient が Flink 1.2 でシェードされない](#)」の問題に対応しました。

- Presto

- LDAP 認証のサポートが追加されました。Amazon EMR の Presto で LDAP を使用する場合は、Presto コーディネーターの HTTPS アクセスを有効にする必要があります (config.properties の http-server.https.enabled=true)。設定の詳細については、Presto ドキュメントの「[LDAP Authentication](#)」を参照してください。
- SHOW GRANTS のサポートが追加されました。

- Amazon EMR の基本 Linux AMI

- 現在、Amazon EMR リリースは Amazon Linux 2017.03 に基づいています。詳細については、「[Amazon Linux AMI 2017.03 リリースノート](#)」を参照してください。

- Python 2.6 は Amazon EMR の基本 Linux イメージから削除されました。Python 2.7 と 3.4 がデフォルトでインストールされます。必要な場合には Python 2.6 を手動でインストールできます。

### 5.5.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.5.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.16.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。

コンポーネント	Version	説明
flink-client	1.2.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.0	HBase コマンドラインクライアント。
hbase-rest-server	1.3.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.0	HBase に Thrift エンドポイントを提供するサービス。



コンポーネント	Version	説明
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.12.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.170	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.170	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.0	YARN スレーブで必要な Apache Spark ライブラリ。

コンポーネント	Version	説明
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.10	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.10	ZooKeeper コマンドラインクライアント。

## 5.5.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.5.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。

分類	説明
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。

分類	説明
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。

分類	説明
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。

分類	説明
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。



分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.4.1

### 5.4.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.166	0.166	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.7.0	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

### 5.4.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

**⚠ Important**

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.4.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.15.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.2.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-httfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.0	HBase コマンドラインクライアント。
hbase-rest-server	1.3.0	HBase の RESTful HTTP エンドポイントを提供するサービス。

コンポーネント	Version	説明
hbase-thrift-server	1.3.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。

コンポーネント	Version	説明
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.166	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.166	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。



コンポーネント	Version	説明
spark-yarn-slave	2.1.0	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.7.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 5.4.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.4.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。

分類	説明
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。

分類	説明
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。

分類	説明
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.4.0

### 5.4.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.1.4	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0



	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.166	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

## 5.4.0 リリースノート

次のリリースノートには、Amazon EMR 5.4.0 リリースに関する情報が含まれています。Amazon EMR 5.3.0 リリースからの変更が含まれています。

リリース日: 2017 年 3 月 8 日

### アップグレード

- Flink 1.2.0 にアップグレードしました
- HBase 1.3.0 にアップグレードしました
- Phoenix 4.9.0 にアップグレード済み

#### Note

古いバージョンの Amazon EMR から Amazon EMR バージョン 5.4.0 以降にアップグレードして、セカンダリインデックスを使用する場合、[Apache Phoenix のドキュメント](#)で説明されるようにローカルインデックスをアップグレードしてください。Amazon EMR は必要な設定を hbase-site 分類から削除しますが、インデックスはデータを再入力する必要があります。インデックスはオンラインとオフラインでアップグレードできます。オンラインのアップグレードがデフォルトです。これはバージョン4.8.0以降のPhoenixクライアントで初期化する間にインデックスの値が再設定されることを意味します。オフラインアップグレードを指定するには、`phoenix.client.localIndexUpgrade` 設定を `phoenix-site` 分類で `False` に設定してから、SSH をプライマリノードに設定して `psql [zookeeper] -1` を実行します。

- Presto 0.166 にアップグレードしました
- Zeppelin 0.7.0 にアップグレードしました

### 変更と機能強化

- r4 インスタンスへのサポートを追加済み。「[Amazon EC2 インスタンスタイプ](#)」を参照してください。

## 5.4.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は

Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.3.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.15.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.2.0	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
<code>ganglia-monitor</code>	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム

コンポーネント	Version	説明
		アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.3.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.3.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.3.0	HBase コマンドラインクライアント。
hbase-rest-server	1.3.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.3.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-hbase	2.1.1-amzn-0	Hive-hbase クライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server2	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.9.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.9.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.166	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.166	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.7.0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

## 5.4.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.4.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。



分類	説明
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の <code>hadoop-metrics2-phoenix.properties</code> ファイルの値を変更します。
pig-properties	Pig の <code>pig.properties</code> ファイルの値を変更します。
pig-log4j	Pig の <code>log4j.properties</code> ファイルの値を変更します。
presto-log	Presto の <code>log.properties</code> ファイルの値を変更します。
presto-config	Presto の <code>config.properties</code> ファイルの値を変更します。
presto-connector-blackhole	Presto の <code>blackhole.properties</code> ファイルの値を変更します。
presto-connector-cassandra	Presto の <code>cassandra.properties</code> ファイルの値を変更します。
presto-connector-hive	Presto の <code>hive.properties</code> ファイルの値を変更します。
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。
presto-connector-kafka	Presto の <code>kafka.properties</code> ファイルの値を変更します。
presto-connector-localfile	Presto の <code>localfile.properties</code> ファイルの値を変更します。
presto-connector-mongodb	Presto の <code>mongodb.properties</code> ファイルの値を変更します。
presto-connector-mysql	Presto の <code>mysql.properties</code> ファイルの値を変更します。

分類	説明
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。

分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.3.2

### 5.3.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.4	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

### 5.3.2 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。



### 5.3.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.14.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.1.4	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。

コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。

コンポーネント	Version	説明
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 5.3.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.3.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。

分類	説明
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。

分類	説明
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。



分類	説明
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。

分類	説明
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。

分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.3.1

### 5.3.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

### 5.3.1 リリースノート

次のリリースノートには、Amazon EMR 5.3.1 リリースに関する情報が含まれています。Amazon EMR 5.3.0 リリースからの変更が含まれています。

リリース日: 2017 年 2 月 7 日

Zeppelin パッチをバックポートして Amazon EMR のデフォルトの AMI を更新するための小さな変更。

### 5.3.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.14.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>flink-client</code>	1.1.4	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。

コンポーネント	Version	説明
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。



コンポーネント	Version	説明
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。

コンポーネント	Version	説明
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 5.3.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.3.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。

分類	説明
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。

分類	説明
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-parquet-logging	Hive の parquet-logging.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。

分類	説明
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の <code>hadoop-metrics2-phoenix.properties</code> ファイルの値を変更します。
pig-properties	Pig の <code>pig.properties</code> ファイルの値を変更します。
pig-log4j	Pig の <code>log4j.properties</code> ファイルの値を変更します。
presto-log	Presto の <code>log.properties</code> ファイルの値を変更します。
presto-config	Presto の <code>config.properties</code> ファイルの値を変更します。
presto-connector-blackhole	Presto の <code>blackhole.properties</code> ファイルの値を変更します。
presto-connector-cassandra	Presto の <code>cassandra.properties</code> ファイルの値を変更します。
presto-connector-hive	Presto の <code>hive.properties</code> ファイルの値を変更します。
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。
presto-connector-kafka	Presto の <code>kafka.properties</code> ファイルの値を変更します。
presto-connector-localfile	Presto の <code>localfile.properties</code> ファイルの値を変更します。
presto-connector-mongodb	Presto の <code>mongodb.properties</code> ファイルの値を変更します。
presto-connector-mysql	Presto の <code>mysql.properties</code> ファイルの値を変更します。



分類	説明
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。

分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.3.0

### 5.3.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

### 5.3.0 リリースノート

次のリリースノートには、Amazon EMR 5.3.0 リリースに関する情報が含まれています。Amazon EMR 5.2.1 リリースからの変更が含まれています。

リリース日: 2017 年 1 月 26 日

#### アップグレード

- Hive 2.1.1 にアップグレードしました
- Hue 3.11.0 にアップグレードしました
- Spark 2.1.0 にアップグレードしました

- Oozie 4.3.0 にアップグレードしました
- Flink 1.1.4 にアップグレードしました

## 変更と機能強化

- `interpreters_shown_on_wheel` ファイルでの順序にかかわらず、ノートブックの選択ホイールで最初に表示するインタープリタを指定する `hue.ini` 設定を使用できるようにするパッチを Hue に追加しました。
- Hive の `parquet-logging.properties` ファイルで値を設定するために使用できる、`hive-parquet-logging` 設定分類を追加しました。

### 5.3.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.14.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.1.4	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。

コンポーネント	Version	説明
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.1-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.1-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.1-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	2.1.1-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.1-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。



コンポーネント	Version	説明
hue-server	3.11.0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.3.0	Oozie コマンドラインクライアント。
oozie-server	4.3.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.1.0	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.1.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.1.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.1.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 5.3.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.3.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の <code>log4j2.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-beeline-log4j2	Hive の <code>beeline-log4j2.properties</code> ファイルの値を変更します。
hive-parquet-logging	Hive の <code>parquet-logging.properties</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。

分類	説明
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。

分類	説明
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。



## Amazon EMR リリース 5.2.3

### 5.2.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.8

### 5.2.3 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.2.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.13.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.1.3	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリ

コンポーネント	Version	説明
		クスの表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.2	Spark コマンドラインクライアント。
spark-history-server	2.0.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。



コンポーネント	Version	説明
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 5.2.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.2.3 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。

分類	説明
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。

分類	説明
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。

分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。

分類	説明
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.2.2

### 5.2.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.1



	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.8	3.4.8

## 5.2.2 リリースノート

次のリリースノートには、Amazon EMR 5.2.2 リリースに関する情報が含まれています。Amazon EMR 5.2.1 リリースからの変更が含まれています。

リリース日: 2017 年 5 月 2 日

以前のリリースから解決された既知の問題

- [SPARK-194459](#) をしました。char/varchar の列を持つ ORC テーブルからの読み取りが失敗する問題に対応しました。

## 5.2.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオー

プンスソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.13.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.1.3	Apache Flink のコマンドラインクライアントスクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.2	Spark コマンドラインクライアント。
spark-history-server	2.0.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

## 5.2.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.2.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。

分類	説明
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。



分類	説明
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。

分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。

分類	説明
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。

分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.2.1

### 5.2.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.1	2.0.1

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.8	3.4.8	3.4.8

## 5.2.1 リリースノート

次のリリースノートには、Amazon EMR 5.2.1 リリースに関する情報が含まれています。Amazon EMR 5.2.0 リリースからの変更が含まれています。

リリース日: 2016 年 12 月 29 日

### アップグレード

- Presto を 0.157.1 にアップグレードしました。詳細については、Presto ドキュメントの「[Presto リリースノート](#)」を参照してください。
- ZooKeeper を 3.4.9 にアップグレードしました。詳細については、Apache ZooKeeper ドキュメントの[ZooKeeper 「リリースノート」](#)を参照してください。

### 変更と機能強化

- 5.0.0、5.0.3、5.2.0 を除く Amazon EMR バージョン 4.8.3 以降の Amazon EC2 の m4.16xlarge インスタンスタイプのサポートが追加されました。
- 現在、Amazon EMR リリースは Amazon Linux 2016.09 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/>」を参照してください。
- これで、Flink および YARN 設定パスの場所はデフォルトで /etc/default/flink に設定されるため、Flink のジョブを起動するために flink または yarn-session.sh ドライバースクリプト

トを実行するときに、環境変数 `FLINK_CONF_DIR` および `HADOOP_CONF_DIR` を設定する必要はありません。

- `FlinkKinesisConsumer` クラスのサポートを追加しました。

#### 以前のリリースから解決された既知の問題

- Hadoop で、レプリケーションと大きなクラスター内の同じファイルの削除の間に競合が発生するため、`ReplicationMonitor` スレッドが長時間停止する問題を修正しました。
- ジョブステータスが正常に更新されなかったときに `ControlledJob`、`#toString` が `null` ポインタ例外 (NPE) で失敗する問題を修正しました。

### 5.2.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。



コンポーネント	Version	説明
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.13.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.1.3	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。

コンポーネント	Version	説明
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.2	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.0.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 5.2.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.2.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の <code>log4j2.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-beeline-log4j2	Hive の <code>beeline-log4j2.properties</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の <code>hive-exec-log4j2.properties</code> ファイルの値を変更します。

分類	説明
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。



分類	説明
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。

分類	説明
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。

分類	説明
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.2.0

### 5.2.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Span](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.1	2.0.1	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

## 5.2.0 リリースノート

次のリリースノートには、Amazon EMR 5.2.0 リリースに関する情報が含まれています。Amazon EMR 5.1.0 リリースからの変更が含まれています。

リリース日: 2016 年 11 月 21 日

### 変更と機能強化

- HBase の Amazon S3 ストレージモードを追加しました。
- HBase ルートディレクトリの Amazon S3 ロケーションを指定できます。Amazon S3 の HBase の詳細については、「[Amazon S3 の HBase](#)」を参照してください。

### アップグレード

- Spark 2.0.2 にアップグレードしました

### 以前のリリースから解決された既知の問題

- EBS のみのインスタンスタイプで 2 TB に制約されていた /mnt に関する問題を修正。
- 通常の log4j-configured .log ファイルではなく、対応する .out ファイルに出力され、1 時間ごとにローテーションされていた、インスタンスコントローラーおよび logpusher ログに関する問題を修正。 .out ファイルはローテーションしないため、最終的には /emr パーティションがいっぱいになります。この問題は、ハードウェア仮想マシン (HVM) のインスタンスタイプにのみ影響します。

## 5.2.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオー

プンスソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.12.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.1.3	Apache Flink のコマンドラインクライアントスクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。



コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.152.3	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.152.3	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.2	Spark コマンドラインクライアント。
spark-history-server	2.0.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

## 5.2.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.2.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。

分類	説明
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase	Apache HBase 用の Amazon EMR-curated 設定。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。

分類	説明
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。

分類	説明
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。

分類	説明
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。



分類	説明
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.1.1

### 5.1.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
Delta	-	-	-	-
Flink	1.1.3	1.1.3	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.152.3	0.152.3	0.150
Spark	2.0.1	2.0.1	2.0.1	2.0.0

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

### 5.1.1 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

#### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

### 5.1.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオー

プンスソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.11.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.1.3	Apache Flink のコマンドラインクライアントスクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.152.3	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。



コンポーネント	Version	説明
presto-worker	0.152.3	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.1	Spark コマンドラインクライアント。
spark-history-server	2.0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

### 5.1.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.1.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。

分類	説明
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。

分類	説明
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。

分類	説明
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。

分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。

分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.1.0

### 5.1.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Flink](#)、[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Span](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2



	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
HBase	1.2.3	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.152.3	0.150	0.150
Spark	2.0.1	2.0.1	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

## 5.1.0 リリースノート

次のリリースノートには、Amazon EMR 5.1.0 リリースに関する情報が含まれています。Amazon EMR 5.0.3 リリースからの変更が含まれています。

リリース日: 2016 年 11 月 3 日

### 変更と機能強化

- Flink 1.1.3 のサポートを追加。
- Presto が、Hue のノートブックセクションでオプションとして追加。

### アップグレード

- HBase 1.2.3 にアップグレードしました
- Zeppelin 0.6.2 にアップグレードしました

### 以前のリリースから解決された既知の問題

- 以前の Amazon EMR 4.x バージョンと同じく、ORC ファイルがある Amazon S3 の Tez クエリが実行されない問題を修正しました。

## 5.1.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.11.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
flink-client	1.1.3	Apache Flink のコマンドラインクライアント スクリプトとアプリケーション。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。

コンポーネント	Version	説明
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-resourcemanager	2.7.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.3	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.3	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.3	HBase コマンドラインクライアント。
hbase-rest-server	1.2.3	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.3	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.152.3	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.152.3	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.1	Spark コマンドラインクライアント。
spark-history-server	2.0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.6.2	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

### 5.1.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.1.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
flink-conf	flink-conf.yaml の設定を変更します。
flink-log4j	Flink log4j.properties の設定を変更します。
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 設定を変更します。



分類	説明
flink-log4j-cli	Flink log4j-cli.properties の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-props-hive-site	HCatalog の props-hive-site.xml の値を変更します。

分類	説明
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。

分類	説明
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。

分類	説明
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。

分類	説明
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。

分類	説明
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.0.3

### 5.0.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)、[Sqoop](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

### 5.0.3 リリースノート

次のリリースノートには、Amazon EMR 5.0.3 リリースに関する情報が含まれています。Amazon EMR 5.0.0 リリースからの変更が含まれています。

リリース日: 2016 年 10 月 24 日

#### アップグレード

- Hadoop 2.7.3 にアップグレードしました
- Presto 0.152.3 にアップグレードします。このアップグレードには Presto ウェブインターフェイスのサポートが含まれています。Presto コーディネーターの Presto ウェブインターフェイスには、ポート 8889 を使用してアクセスできます。Presto ウェブインターフェイスの詳細については、Presto ドキュメントの「[ウェブインターフェイス](#)」を参照してください。
- Spark 2.0.1 にアップグレードしました
- 現在、Amazon EMR リリースは Amazon Linux 2016.09 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/>」を参照してください。



### 5.0.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.10.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>ganglia-monitor</code>	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム

コンポーネント	Version	説明
		アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.7.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.152.3	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.152.3	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.1	Spark コマンドラインクライアント。
spark-history-server	2.0.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

### 5.0.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.0.3 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。

分類	説明
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。

分類	説明
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。



分類	説明
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。

分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。

分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.0.2

### 5.0.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)、[Sqoop](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

## 5.0.2 リリースノート

### 5.0.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	4.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。

コンポーネント	Version	説明
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.9.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.2-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-httfs-server	2.7.2-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。



コンポーネント	Version	説明
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。

コンポーネント	Version	説明
mysql-server	5.5.46	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.150	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.150	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.0	Spark コマンドラインクライアント。
spark-history-server	2.0.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.0	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	2.0.0	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.1-SNAPSHOT	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

## 5.0.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-5.0.2 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。

分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。

分類	説明
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。

分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。

分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。



分類	説明
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.0.1

### 5.0.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)、[Sqoop](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

## 5.0.1 リリースノート

### 5.0.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	4.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.9.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	2.7.2-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.46	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.150	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.150	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.0	Spark コマンドラインクライアント。
spark-history-server	2.0.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.1-SNAPSHOT	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。



コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

### 5.0.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-5.0.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の <code>log4j2.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-beeline-log4j2	Hive の <code>beeline-log4j2.properties</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の <code>hive-exec-log4j2.properties</code> ファイルの値を変更します。

分類	説明
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。

分類	説明
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。

分類	説明
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します。

分類	説明
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 5.0.0

### 5.0.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie](#)、[Phoenix](#)、[Pig](#)、[Presto](#)、[Spark](#)、[Sqoop](#)

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8



## 5.0.0 リリースノート

リリース日: 2016 年 7 月 27 日

### アップグレード

- Hive 2.1 にアップグレードしました
- Presto 0.150 にアップグレードしました
- Spark 2.0 にアップグレードしました
- Hue 3.10.0 にアップグレードしました
- Pig 0.16.0 にアップグレードしました
- Tez 0.8.4 にアップグレードしました
- Zeppelin 0.6.1 にアップグレードしました

### 変更と機能強化

- Amazon EMR は Hive (バージョン 2.1)、Pig (バージョン 0.16.0) の最新のオープンソースバージョンをサポートします。以前に Amazon EMR で Hive または Pig を使用している場合、このことはいくつかのユースケースに影響を与える可能性があります。詳細については、「[Hive](#)」および「[Pig](#)」を参照してください。
- 現在の Hive および Pig のデフォルト実行エンジンは Tez です。これを変更するには、hive-site および pig-properties のそれぞれの設定分類の適切な値を編集します。
- サービスが原因を識別できる場合にステップ障害の根本原因を表示できるようにするステップ、デバッグ機能が追加されました。詳細については、「Amazon EMR 管理ガイド」の「[デバッグの詳細ステップ](#)」を参照してください。
- アプリケーションは以前「-Sandbox」で終了しましたが、そのサフィックスはもうありません。たとえば、これらのアプリケーションを使ってクラスターを起動するスクリプトを使用している場合、これによってオートメーションが中断する可能性があります。次の表は、Amazon EMR 4.7.2 と Amazon EMR 5.0.0 のアプリケーション名を示しています。

### アプリケーション名変更

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-サンドボックス	ZooKeeper

- Spark は現在 Scala 2.11 向けにコンパイルされています。
- 現在のデフォルト JVM は Java 8 です。すべてのアプリケーションは Java 8 ランタイムを使用して動作します。アプリケーションのバイトコードターゲットには変更はありません。ほとんどのアプリケーションは、引き続き Java 7 を対象としています。
- Zeppelin には、認証機能が組み込まれています。詳細については、「[Zeppelin](#)」を参照してください。
- セキュリティ設定のサポートを追加しました。これにより、暗号化オプションをより簡単に作成、適用できます。詳細については、「[Data Encryption](#)」を参照してください。

## 5.0.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。

コンポーネント	Version	説明
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.9.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.2-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.7.2-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。

コンポーネント	Version	説明
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	2.1.0-amzn-0	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	2.1.0-amzn-0	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	2.1.0-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	2.1.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	2.1.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	3.10.0-amzn-0	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.46	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.150	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.150	いくつかのクエリを実行するサービス。
pig-client	0.16.0-amzn-0	Pig コマンドラインクライアント。
spark-client	2.0.0	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	2.0.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	2.0.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	2.0.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.1-SNAPSHOT	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

## 5.0.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-5.0.0 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。



分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j2	HCatalog WebHCat の log4j2.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-beeline-log4j2	Hive の beeline-log4j2.properties ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j2	Hive の hive-exec-log4j2.properties ファイルの値を変更します。
hive-llap-daemon-log4j2	Hive の llap-daemon-log4j2.properties ファイルの値を変更します。
hive-log4j2	Hive の hive-log4j2.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。

分類	説明
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。

分類	説明
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-hive-site	Spark の hive-site.xml ファイルの値を変更します
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。

分類	説明
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR 4.x リリースバージョン

このセクションでは、各 Amazon EMR 4.x リリースバージョンのアプリケーションバージョン、リリースノート、コンポーネントバージョン、および設定分類を示しています。

クラスターを起動するときに、複数の Amazon EMR リリースから選択できます。これにより、互換性の要件に適合するアプリケーションバージョンをテストして使用できます。リリース番号はリリースラベルで指定します。リリースラベルの形式は `emr-x.x.x` です。例えば `emr-7.1.0` です。

新しい Amazon EMR リリースは、特定のリージョンで最初に利用可能になり、その後数日で、他のリージョンでも利用できるようになります。この期間中、お客様のリージョンで最新のリリースバージョンが利用できない場合があります。

すべての Amazon EMR 4.x リリースのアプリケーションバージョンの包括的な表については、「[Amazon EMR 4.x リリースのアプリケーションバージョン](#)」を参照してください。

### トピック

- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースバージョンの違い](#)
- [Amazon EMR リリース 4.9.6](#)

- [Amazon EMR リリース 4.9.5](#)
- [Amazon EMR リリース 4.9.4](#)
- [Amazon EMR リリース 4.9.3](#)
- [Amazon EMR リリース 4.9.2](#)
- [Amazon EMR リリース 4.9.1](#)
- [Amazon EMR リリース 4.8.5](#)
- [Amazon EMR リリース 4.8.4](#)
- [Amazon EMR リリース 4.8.3](#)
- [Amazon EMR リリース 4.8.2](#)
- [Amazon EMR リリース 4.8.1](#)
- [Amazon EMR リリース 4.8.0](#)
- [Amazon EMR リリース 4.7.4](#)
- [Amazon EMR リリース 4.7.3](#)
- [Amazon EMR リリース 4.7.2](#)
- [Amazon EMR リリース 4.7.1](#)
- [Amazon EMR リリース 4.7.0](#)
- [Amazon EMR リリース 4.6.1](#)
- [Amazon EMR リリース 4.6.0](#)
- [Amazon EMR リリース 4.5.0](#)
- [Amazon EMR リリース 4.4.0](#)
- [Amazon EMR リリース 4.3.0](#)
- [Amazon EMR リリース 4.2.0](#)
- [Amazon EMR リリース 4.1.0](#)
- [Amazon EMR リリース 4.0.0](#)

## Amazon EMR 4.x リリースのアプリケーションバージョン

各 Amazon EMR 4.x リリースで利用可能なアプリケーションバージョンを示した包括的な表については、ブラウザで「[Amazon EMR 4.x リリースのアプリケーションバージョン](#)」を開いて確認してください。

## Amazon EMR 4.x リリースバージョンの違い

「Amazon EMR 管理ガイド」の Amazon EMR 機能のドキュメントでは、特定の機能が使用可能になった Amazon EMR リリースバージョンを示すとともに、Amazon EMR の機能の違いを 4.0.0 にまでさかのぼって示しています。

Amazon EMR リリース 5.0.0 バージョン以降、一部のアプリケーションでは大幅なバージョンアップグレードが行われ、インストールまたは操作の詳細が変更されました。また、サンドボックスアプリケーションからネイティブアプリケーションに昇格されたものもあります。このセクションの各トピックでは、Amazon EMR 4.x リリースバージョンを使用する際のアプリケーション特有の大きな違いを説明しています。

### トピック

- [サンドボックスアプリケーション](#)
- [Amazon EMR 4.x での Hive の使用に関する考慮事項](#)
- [Amazon EMR 4.x での Pig の使用に関する考慮事項](#)

### サンドボックスアプリケーション

Amazon EMR 4.x リリースバージョンを使用する場合、一部のアプリケーションはサンドボックスアプリケーションとみなされます。サンドボックスアプリケーションは、アプリケーションの初期バージョンで、需要があったために最初の Amazon EMR リリースの時点で使用可能になりました。コンソール、AWS CLI、または API を使用して、Amazon EMR でサンドボックスアプリケーションをネイティブアプリケーションと同じようにインストールできますが、サンドボックスアプリケーションはサポートとドキュメントが限られています。サンドボックスアプリケーションは、Amazon EMR リリースバージョン 5.0.0 以降にネイティブの完全にサポートされたアプリケーションとなりました。以下は、Amazon EMR 4.x リリースバージョンのサンドボックスアプリケーションです。

- Oozie
- Presto
- Sqoop
- Zeppelin
- ZooKeeper

サンドボックスアプリケーションをインストールするときは、アプリケーション名にサフィックス `-sandbox` を付けます。たとえば、サンドボックスバージョンの *Presto* をインストールするに

は、Presto-sandbox を使用します。インストールは、完全にサポートされているアプリケーションの場合よりも時間がかかります。このセクションで、各アプリケーションのバージョン番号のリストは、コミュニティバージョンのアプリケーションに対応しています。

### Oozie (サンドボックスバージョン)

Oozie は、Amazon EMR リリースバージョン 4.1.0 からサンドボックスアプリケーションとして利用可能になりました。

Oozie のサンプルは、サンドボックスを使用したデフォルトではインストールされません。サンプルをインストールするため、マスターノードに SSH 接続し、`install-oozie-examples` を実行します。

### Oozie-Sandbox のバージョン情報

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.9.6	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.9.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager



Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
		cemanager, oozie-client, oozie-server
emr-4.9.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcecemanager, oozie-client, oozie-server
emr-4.9.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcecemanager, oozie-client, oozie-server

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.9.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.9.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.8.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.8.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.8.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.8.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.8.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.8.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.7.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.7.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.7.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.7.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.7.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.6.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server
emr-4.6.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, oozie-client, oozie-server



Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.5.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.4.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.3.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR リリースラベル	Oozie-Sandbox のバージョン	Oozie-Sandbox でインストールされるコンポーネント
emr-4.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.1.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

### Presto (サンドボックスバージョン)

Presto は、Amazon EMR リリースバージョン 4.1.0 からサンドボックスアプリケーションとして利用可能になりました。

### Presto-Sandbox のバージョン情報

Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
emr-4.9.6	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
		hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
emr-4.9.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
emr-4.8.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
emr-4.8.2	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.1	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.0	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
emr-4.7.4	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.3	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.2	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
emr-4.7.1	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.0	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.6.1	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker



Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
emr-4.6.0	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.5.0	0.140	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.4.0	0.136	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto-Sandbox のバージョン	Presto-Sandbox でインストールされるコンポーネント
emr-4.3.0	0.130	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.2.0	0.125	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.1.0	0.119	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

### Sqoop (サンドボックスバージョン)

Sqoop は、Amazon EMR リリースバージョン 4.4.0 からサンドボックスアプリケーションとして利用可能になりました。

## Sqoop-Sandbox のバージョン情報

Amazon EMR リリースラベル	Sqoop-Sandbox のバージョン	Sqoop-Sandbox でインストールされるコンポーネント
emr-4.9.6	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop-Sandbox のバージョン	Sqoop-Sandbox でインストールされるコンポーネント
emr-4.9.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop-Sandbox のバージョン	Sqoop-Sandbox でインストールされるコンポーネント
emr-4.8.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop-Sandbox のバージョン	Sqoop-Sandbox でインストールされるコンポーネント
emr-4.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop-Sandbox のバージョン	Sqoop-Sandbox でインストールされるコンポーネント
emr-4.7.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop-Sandbox のバージョン	Sqoop-Sandbox でインストールされるコンポーネント
emr-4.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client



Amazon EMR リリースラベル	Sqoop-Sandbox のバージョン	Sqoop-Sandbox でインストールされるコンポーネント
emr-4.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client
emr-4.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client

### Zeppelin (サンドボックスバージョン)

Zeppelin は、Amazon EMR リリースバージョン 4.1.0 からサンドボックスアプリケーションとして利用可能になりました。

## Zeppelin-Sandbox のバージョン情報

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.9.6	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
		on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.9.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.8.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.8.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.8.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.4	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.7.3	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.2	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server



Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.7.1	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.6.1	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.6.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.5.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.4.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin-Sandbox のバージョン	Zeppelin-Sandbox でインストールされるコンポーネント
emr-4.3.0	0.5.5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resource manager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.2.0	0.5.5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resource manager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.1.0	0.6.0-SNAPSHOT	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resource manager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

## Zookeeper (サンドボックスバージョン)

Zookeeper は、Amazon EMR リリースバージョン 4.6.0 からサンドボックスアプリケーションとして利用可能になりました。

### ZooKeeper-Sandbox のバージョン情報

Amazon EMR リリースラベル	ZooKeeper-Sandbox のバージョン	ZooKeeper-Sandbox でインストールされるコンポーネント
emr-4.9.6	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, zookeeper-client, zookeeper-server
emr-4.9.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, zookeeper-client, zookeeper-server
emr-4.9.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager

Amazon EMR リリースラベル	ZooKeeper-Sandbox のバージョン	ZooKeeper-Sandbox でインストールされるコンポーネント
		cemanager, zookeeper-client, zookeeper-server
emr-4.9.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, zookeeper-client, zookeeper-server
emr-4.9.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, zookeeper-client, zookeeper-server
emr-4.9.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper-Sandbox のバージョン	ZooKeeper-Sandbox でインストールされるコンポーネント
emr-4.8.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, zookeeper-client, zookeeper-server
emr-4.8.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, zookeeper-client, zookeeper-server
emr-4.8.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper-Sandbox のバージョン	ZooKeeper-Sandbox でインストールされるコンポーネント
emr-4.8.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper-Sandbox のバージョン	ZooKeeper-Sandbox でインストールされるコンポーネント
emr-4.7.4	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	ZooKeeper-Sandbox のバージョン	ZooKeeper-Sandbox でインストールされるコンポーネント
emr-4.7.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.6.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper-Sandbox のバージョン	ZooKeeper-Sandbox でインストールされるコンポーネント
emr-4.6.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

## Amazon EMR 4.x での Hive の使用に関する考慮事項

このセクションでは、Amazon EMR 4.x リリースバージョンで Hive バージョン 1.0.0 を使用する場合と Amazon EMR 5.x リリースバージョンで Hive 2.x を使用する場合に考慮すべき違いについて説明します。

### ACID トランザクションはサポート外

Amazon EMR 4.x リリースバージョンでの Hive は、4.x リリースバージョンの使用時に Amazon S3 に保存された Hive データとの ACID トランザクションをサポートしていません。Amazon S3 でトランザクションテーブルを作成しようとする、例外が発生します。

### Amazon S3 でのテーブルへの読み取りと書き込み

Amazon EMR 4.x リリースバージョンでの Hive は、一時ファイルを使用せずに直接 Amazon S3 に書き込むことができます。これにより、パフォーマンスは向上しますが、結果として Amazon S3 の同じテーブルに同じ Hive ステートメントで読み取りと書き込みを行うことができません。回避策は、HDFS に一時テーブルを作成して使用することです。

次の例は、複数の Hive ステートメントを使用して Amazon S3 内のテーブルを更新する方法を示しています。ステートメントは、Amazon S3 にある `my_s3_table` という名前のテーブルに基づいて、HDFS に `tmp` という名前の一時テーブルを作成します。これにより、Amazon S3 のテーブルは一時テーブルの内容で更新されます。

```
CREATE TEMPORARY TABLE tmp LIKE my_s3_table;
```

```
INSERT OVERWRITE TABLE tmp SELECT ....;  
INSERT OVERWRITE TABLE my_s3_table SELECT * FROM tmp;
```

## Log4j vs. Log4j 2

Amazon EMR 4.x リリースバージョンでの Hive は Log4j を使用します。バージョン 5.0.0 以降では、Log4j 2 がデフォルトです。これらのバージョンでは、異なるログ設定が必要になる場合があります。詳細については、[Apache Log4j 2](#) を参照してください。

## MapReduce がデフォルトの実行エンジン

Amazon EMR 4.x リリースバージョンでの Hive では、デフォルトの実行エンジンとして MapReduce を使用します。Amazon EMR バージョン 5.0.0 以降では、Tez がデフォルトです。これにより、ほとんどのワークフローでパフォーマンスが向上します。

## Hive 認可

Amazon EMR 4.x リリースバージョンでの Hive では、HDFS の [Hive 認可](#) をサポートしますが、EMRFS および Amazon S3 はサポートしません。Amazon EMR クラスターは、デフォルトで認可が無効な状態で実行されます。

## Amazon S3 での Hive ファイルマージ動作

Amazon EMR 4.x リリースバージョンでの Hive は、hive.merge.mapfiles が true の場合、マップのみのジョブの最後で小さいファイルをマージします。マージは、ジョブの平均出力サイズが hive.merge.smallfiles.avgsize の設定未満の場合にのみトリガーされます。最終出力パスが HDFS 内にある場合は、Amazon EMR Hive でも動作はまったく同じです。ただし、出力パスが Amazon S3 にある場合、hive.merge.smallfiles.avgsize パラメータは無視されます。その場合、hive.merge.mapfiles が true に設定されているとマージタスクが常にトリガされます。

## Amazon EMR 4.x での Pig の使用に関する考慮事項

Pig バージョン 0.14.0 は、Amazon EMR 4.x リリースバージョンを使用して作成されたクラスターにインストールされます。Pig は、Amazon EMR 5.0.0 でバージョン 0.16.0 にアップグレードされました。大きな違いは以下で説明します。

### デフォルト実行エンジンの相違

Amazon EMR 4.x リリースバージョンでの Pig バージョン 0.14.0 では、デフォルトの実行エンジンとして MapReduce を使用します。Pig 0.16.0 以降では、Apache Tez を使用しま

す。exectype=mapreduce 設定分類で明示的に pig-properties と設定すると、MapReduce を使用できます。

### 削除された Pig ユーザ定義関数 (UDF)

Amazon EMR 4.x リリースバージョンでの Pig で使用可能だったカスタム UDF は、Pig 0.16.0 以降は削除されました。ほとんどの UDF には代わりに使用できる同等の関数があります。削除された UDF と同等の関数のリストを次に示します。詳細については、Apache Pig サイトの「[組み込み関数](#)」を参照してください。

削除された UDF	同等の関数
FORMAT_DT(dtformat, date)	GetHour(date), GetMinute(date), GetMonth(date), GetSecond(date), GetWeek(date), GetYear(date), GetDay(date)
EXTRACT(string, pattern)	REGEX_EXTRACT_ALL(string, pattern)
REPLACE(string, pattern, replacement)	REPLACE(string, pattern, replacement)
DATE_TIME()	ToDate()
DURATION(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween(dt, dt2), SecondsBetween(dt, dt2), MonthsBetween(dt, dt2), MinutesBetween(dt, dt2), HoursBetween(dt, dt2)
EXTRACT_DT(format, date)	GetHour(date), GetMinute(date), GetMonth(date), GetSecond(date), GetWeek(date), GetYear(date), GetDay(date)
OFFSET_DT(date, duration)	AddDuration(date, duration), SubtractDuration(date, duration)
PERIOD(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween(dt, dt2), SecondsBetween(dt, dt2), MonthsBetween(dt, dt2), MinutesBetween(dt, dt2), HoursBetween(dt, dt2)
CAPITALIZE(string)	UCFIRST(string)

削除された UDF	同等の関数
CONCAT_WITH()	CONCAT()
INDEX_OF()	INDEXOF()
LAST_INDEX_OF()	LAST_INDEXOF()
SPLIT_ON_REGEX()	STRSPLT()
UNCAPITALIZE()	LCFIRST()

次の UDF は削除され、同等物はありません。

FORMAT()、LOCAL\_DATE()、LOCAL\_TIME()、CENTER()、LEFT\_PAD()、REPEAT()、REPLACE\_

使用停止になった Grunt コマンド

Pig 0.16.0 以降ではいくつかの Grunt コマンドが使用停止されました。次の表に、Pig 0.14.0 の Grunt コマンドおよび最新バージョンで適用可能な同等のコマンドを示します。

Pig 0.14.0 と、同等の現在の Grunt コマンド

Pig 0.14.0 Grunt コマンド	0.16.0 以降の Pig Grunt コマンド
cat <non-hdfs-path>)	fs -cat <non-hdfs-path>;
cd <non-hdfs-path>;	同等物なし
ls <non-hdfs-path>;	fs -ls <non-hdfs-path>;
move <non-hdfs-path> <non-hdfs-path>;	fs -mv <non-hdfs-path> <non-hdfs-path>;
copy <non-hdfs-path> <non-hdfs-path>;	fs -cp <non-hdfs-path> <non-hdfs-path>;
copyToLocal <non-hdfs-path> <local-path>;	fs -copyToLocal <non-hdfs-path> <local-path>;
copyFromLocal <local-path> <non-hdfs-path>;	fs -copyFromLocal <local-path> <non-hdfs-path>;
mkdir <non-hdfs-path>;	fs -mkdir <non-hdfs-path>;

Pig 0.14.0 Grunt コマンド	0.16.0 以降の Pig Grunt コマンド
<code>rm &lt;non-hdfs-path&gt;;</code>	<code>fs -rm -r -skipTrash &lt;non-hdfs-path&gt;;</code>
<code>rmf &lt;non-hdfs-path&gt;;</code>	<code>fs -rm -r -skipTrash &lt;non-hdfs-path&gt;;</code>

## Non-HDFS ホームディレクトリについて削除された機能

Amazon EMR 4.x リリースバージョンの Pig 0.14.0 には、hadoop ユーザー以外のホームディレクトリを持たないユーザーが Pig スクリプトを実行できるようにする 2 つのメカニズムがあります。第 1 のメカニズムは、自動フォールバックで、ホームディレクトリが存在しない場合、ルートディレクトリに最初の作業ディレクトリを設定します。2 つめは、最初の作業ディレクトリを変更する `pig.initial.fs.name` プロパティです。

Amazon EMR 5.0.0 以降のバージョンでは、これらのメカニズムを使用できないので、ユーザーは HDFS 上のホームディレクトリが必要です。これは、hadoop ユーザーには当てはまりません。起動時にホームディレクトリがプロビジョニングされるからです。 `command-runner.jar` を使用して別のユーザーが明示的に指定されていない限り、スクリプトは、Hadoop ユーザーにデフォルトの Hadoop jar ステップを使用して動作します。

## Amazon EMR リリース 4.9.6

### 4.9.6 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-



	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

## 4.9.6 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

## 4.9.6 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。

コンポーネント	Version	説明
emr-goodies	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.17.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。

コンポーネント	Version	説明
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-9	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-9	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-9	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-9	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-9	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

#### 4.9.6 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-4.9.6 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。



分類	説明
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。

分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。

分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。

分類	説明
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.9.5

### 4.9.5 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

## 4.9.5 リリースノート

次のリリースノートには、Amazon EMR リリース 4.9.5 に関する情報が含まれています。4.9.4 からの変更が含まれています。

初回リリース日: 2018 年 8 月 29 日

変更、機能強化、解決した問題

- HBase

- このリリースでは、潜在的なセキュリティ脆弱性に対処しています。

#### 4.9.5 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.17.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。



コンポーネント	Version	説明
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	1.0.0-amzn-9	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-9	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-9	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-9	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-9	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。

コンポーネント	Version	説明
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

#### 4.9.5 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.9.5 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。

分類	説明
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。

分類	説明
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。

分類	説明
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。

分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。



分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.9.4

### 4.9.4 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

#### 4.9.4 リリースノート

次のリリースノートには、Amazon EMR リリース 4.9.4 に関する情報が含まれています。4.9.3 からの変更が含まれています。

初回リリース日: 2018 年 3 月 29 日

##### 変更、機能強化、解決した問題

- 潜在的な脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。

#### 4.9.4 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.17.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>ganglia-monitor</code>	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム

コンポーネント	Version	説明
		アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-9	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-9	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-9	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-9	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-9	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ



コンポーネント	Version	説明
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

#### 4.9.4 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.9.4 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。

分類	説明
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。

分類	説明
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。

分類	説明
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.9.3

### 4.9.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3



	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

### 4.9.3 リリースノート

次のリリースノートには、Amazon EMR 4.9.3 リリースに関する情報が含まれています。Amazon EMR 4.9.2 リリースからの変更が含まれています。

初回リリース日: 2018 年 1 月 22 日

#### 変更、機能強化、解決した問題

- 投機的実行 (CVE-2017-5715、CVE-2017-5753、CVE-2017-5754) に関連する脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。詳細については、「<https://aws.amazon.com/security/security-bulletins/AWS-2018-013/>」を参照してください。

### 4.9.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.17.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリ

コンポーネント	Version	説明
		クスの表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-9	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-9	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-9	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	1.0.0-amzn-9	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-9	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 4.9.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-4.9.3 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。



分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の <code>log4j.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の <code>hive-exec-log4j.properties</code> ファイルの値を変更します。
hive-log4j	Hive の <code>hive-log4j.properties</code> ファイルの値を変更します。

分類	説明
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。

分類	説明
phoenix-hbase-metrics	Phoenix の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
phoenix-hbase-site	Phoenix の <code>hbase-site.xml</code> ファイルの値を変更します。
phoenix-log4j	Phoenix の <code>log4j.properties</code> ファイルの値を変更します。
phoenix-metrics	Phoenix の <code>hadoop-metrics2-phoenix.properties</code> ファイルの値を変更します。
pig-properties	Pig の <code>pig.properties</code> ファイルの値を変更します。
pig-log4j	Pig の <code>log4j.properties</code> ファイルの値を変更します。
presto-log	Presto の <code>log.properties</code> ファイルの値を変更します。
presto-config	Presto の <code>config.properties</code> ファイルの値を変更します。
presto-connector-blackhole	Presto の <code>blackhole.properties</code> ファイルの値を変更します。
presto-connector-cassandra	Presto の <code>cassandra.properties</code> ファイルの値を変更します。
presto-connector-hive	Presto の <code>hive.properties</code> ファイルの値を変更します。
presto-connector-jmx	Presto の <code>jmx.properties</code> ファイルの値を変更します。
presto-connector-kafka	Presto の <code>kafka.properties</code> ファイルの値を変更します。

分類	説明
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.9.2

### 4.9.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

## 4.9.2 リリースノート

次のリリースノートには、Amazon EMR 4.9.2 リリースに関する情報が含まれています。Amazon EMR 4.9.1 リリースからの変更が含まれています。

リリース日: 2017 年 7 月 13 日

このリリースでは小さな変更、バグ修正、および機能強化が行われました。

### 4.9.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.3.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
emr-goodies	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。



コンポーネント	Version	説明
emr-kinesis	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.17.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-9	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-9	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-9	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-9	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-9	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション

コンポーネント	Version	説明
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

## 4.9.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-4.9.2 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。

分類	説明
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。



分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。

分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。

分類	説明
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.9.1

### 4.9.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

## 4.9.1 リリースノート

次のリリースノートには、Amazon EMR 4.9.1 リリースに関する情報が含まれています。Amazon EMR 4.8.4 リリースからの変更が含まれています。

リリース日: 2017 年 4 月 10 日

以前のリリースから解決された既知の問題

- [HIVE-9976](#) および [HIVE-10106](#) を移植しました。
- YARN で、多数のノード (2,000 より大きい) とコンテナ (5,000 より大きい) が out-of-memory エラーを引き起こす問題を修正しました。例: "Exception in thread main java.lang.OutOfMemoryError"。

変更と機能強化

- 現在、Amazon EMR リリースは Amazon Linux 2017.03 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/>」を参照してください。
- Python 2.6 は Amazon EMR の基本 Linux イメージから削除されました。必要な場合には Python 2.6 を手動でインストールできます。

#### 4.9.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。

コンポーネント	Version	説明
emr-goodies	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.3.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.15.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.7.3-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。



コンポーネント	Version	説明
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-9	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-9	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-9	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-9	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-9	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。

コンポーネント	Version	説明
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 4.9.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-4.9.1 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。

分類	説明
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。

分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。

分類	説明
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。



分類	説明
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.8.5

### 4.8.5 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.152.3
Spark	1.6.3	1.6.3	1.6.3	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.8

## 4.8.5 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

**⚠ Important**

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

#### 4.8.5 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.14.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。

コンポーネント	Version	説明
hcatalog-client	1.0.0-amzn-8	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-8	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-8	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-8	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-8	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-8	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。

コンポーネント	Version	説明
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。



コンポーネント	Version	説明
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

#### 4.8.5 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.8.5 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。

分類	説明
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。

分類	説明
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。

分類	説明
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。

分類	説明
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。

分類	説明
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.8.4

### 4.8.4 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.152.3	0.151



	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
Spark	1.6.3	1.6.3	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.8	3.4.8

#### 4.8.4 リリースノート

次のリリースノートには、Amazon EMR 4.8.4 リリースに関する情報が含まれています。Amazon EMR 4.8.3 リリースからの変更が含まれています。

リリース日: 2017 年 2 月 7 日

このリリースでは小さな変更、バグ修正、および機能強化が行われました。

#### 4.8.4 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.14.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-8	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-8	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-8	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-8	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	1.0.0-amzn-8	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-8	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.54+	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.25+	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

#### 4.8.4 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.8.4 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-props-hive-site	HCatalog の <code>props-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の <code>log4j.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の <code>hive-exec-log4j.properties</code> ファイルの値を変更します。
hive-log4j	Hive の <code>hive-log4j.properties</code> ファイルの値を変更します。



分類	説明
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。

分類	説明
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。

分類	説明
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.8.3

### 4.8.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.152.3	0.151	0.151
Spark	1.6.3	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.8	3.4.8	3.4.8

### 4.8.3 リリースノート

次のリリースノートには、Amazon EMR 4.8.3 リリースに関する情報が含まれています。Amazon EMR 4.8.2 リリースからの変更が含まれています。

リリース日: 2016 年 12 月 29 日

#### アップグレード

- Presto を 0.157.1 にアップグレードしました。詳細については、Presto ドキュメントの「[Presto リリースノート](#)」を参照してください。
- Spark を 1.6.3 にアップグレードしました。詳細については、Apache Spark ドキュメントの「[Spark リリースノート](#)」を参照してください。
- を ZooKeeper 3.4.9 にアップグレードしました。詳細については、Apache ZooKeeper ドキュメントの[ZooKeeper 「リリースノート」](#)を参照してください。

#### 変更と機能強化

- 5.0.0、5.0.3、5.2.0 を除く Amazon EMR バージョン 4.8.3 以降の Amazon EC2 の m4.16xlarge インスタンスタイプのサポートが追加されました。
- 現在、Amazon EMR リリースは Amazon Linux 2016.09 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/>」を参照してください。

#### 以前のリリースから解決された既知の問題

- Hadoop で、レプリケーションと大きなクラスター内の同じファイルの削除の間に競合があるため、ReplicationMonitor スレッドが長時間停止する問題を修正しました。
- ジョブステータスが正常に更新されなかったときに ControlledJob、#toString が null ポインタ例外 (NPE) で失敗する問題を修正しました。

### 4.8.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	4.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.2.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.13.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>ganglia-monitor</code>	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステム



コンポーネント	Version	説明
		アプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.3-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.7.3-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.3-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-8	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-8	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。

コンポーネント	Version	説明
hcatalog-webhcat-server	1.0.0-amzn-8	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-8	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-8	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-8	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ

コンポーネント	Version	説明
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.157.1	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.157.1	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.3	Spark コマンドラインクライアント。
spark-history-server	1.6.3	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.3	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.3	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.9	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.9	ZooKeeper コマンドラインクライアント。

### 4.8.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-4.8.3 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。

分類	説明
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。

分類	説明
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。



分類	説明
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。

分類	説明
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.8.2

### 4.8.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.152.3	0.151	0.151	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.3

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

## 4.8.2 リリースノート

次のリリースノートには、Amazon EMR 4.8.2 リリースに関する情報が含まれています。Amazon EMR 4.8.0 リリースからの変更が含まれています。

リリース日: 2016 年 10 月 24 日

### アップグレード

- Hadoop 2.7.3 にアップグレードしました
- Presto 0.152.3 にアップグレードします。このアップグレードには Presto ウェブインターフェイスのサポートが含まれています。Presto コーディネーターの Presto ウェブインターフェイスには、ポート 8889 を使用してアクセスできます。Presto ウェブインターフェイスの詳細については、Presto ドキュメントの「[ウェブインターフェイス](#)」を参照してください。
- 現在、Amazon EMR リリースは Amazon Linux 2016.09 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/>」を参照してください。

## 4.8.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	4.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.10.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.3-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.3-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.3-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.3-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-https-server	2.7.3-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.3-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.3-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.3-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resource-manager	2.7.3-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-7	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-7	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-7	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-7	Hive コマンドラインクライアント。



コンポーネント	Version	説明
hive-metastore-server	1.0.0-amzn-7	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-7	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.52	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.152.3	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.152.3	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.2	Spark コマンドラインクライアント。
spark-history-server	1.6.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

## 4.8.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-4.8.2 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の <code>log4j.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の <code>hive-exec-log4j.properties</code> ファイルの値を変更します。
hive-log4j	Hive の <code>hive-log4j.properties</code> ファイルの値を変更します。

分類	説明
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。

分類	説明
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。

分類	説明
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.8.1

### 4.8.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。



- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.151	0.151	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.5.6	0.5.6

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

## 4.8.1 リリースノート

### 4.8.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	3.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。

コンポーネント	Version	説明
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.9.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.2-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。

コンポーネント	Version	説明
hadoop-httfs-server	2.7.2-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。

コンポーネント	Version	説明
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-7	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-7	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-7	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-7	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-7	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-7	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。

コンポーネント	Version	説明
mysql-server	5.5.51	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.151	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.151	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.2	Spark コマンドラインクライアント。
spark-history-server	1.6.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.2	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	1.6.2	YARN スレーブに必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

#### 4.8.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.8.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。



分類	説明
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。

分類	説明
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。

分類	説明
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。

分類	説明
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。

分類	説明
presto-connector-tpch	Presto の tpcj.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeper の zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.8.0

### 4.8.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.151	0.148	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

## 4.8.0 リリースノート

次のリリースノートには、Amazon EMR 4.8.0 リリースに関する情報が含まれています。Amazon EMR 4.7.2 リリースからの変更が含まれています。

リリース日: 2016 年 9 月 7 日



## アップグレード

- HBase 1.2.2 にアップグレードしました
- Presto-Sandbox 0.151 にアップグレードしました
- Tez 0.8.4 にアップグレードしました
- Zeppelin-Sandbox 0.6.1 にアップグレードしました

## 変更と機能強化

- YARN で、インスタンスが終了したために がもう存在しないコンテナをクリーンアップ ApplicationMaster しようとする問題を修正しました。
- Oozie の例の Hive2 アクションの hive-server2 URL を修正しました。
- さらに多くの Presto カタログのサポートを追加しました。
- 次のパッチを移植しました:  
[HIVE-8948](#)、[HIVE-12679](#)、[HIVE-13405](#)、[PHOENIX-3116](#)、[HADOOP-12689](#)
- セキュリティ設定のサポートを追加しました。これにより、暗号化オプションをより簡単に作成、適用できます。詳細については、「[Data Encryption](#)」を参照してください。

## 4.8.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	3.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.9.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	2.7.2-amzn-4	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-4	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-4	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-4	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-4	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-4	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-4	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-4	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-4	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-4	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.2.2	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.2	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.2	HBase コマンドラインクライアント。
hbase-rest-server	1.2.2	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.2	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-7	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-7	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-7	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-7	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	1.0.0-amzn-7	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-7	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.51	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.151	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.151	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.2	Spark コマンドラインクライアント。
spark-history-server	1.6.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.4	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.6.1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

#### 4.8.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.8.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-props-hive-site	HCatalog の <code>props-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の <code>log4j.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の <code>hive-exec-log4j.properties</code> ファイルの値を変更します。
hive-log4j	Hive の <code>hive-log4j.properties</code> ファイルの値を変更します。



分類	説明
hive-site	Hive の hive-site.xml ファイルの値を変更します
hiveserver2-site	Hive Server2 の hiveserver2-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。

分類	説明
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-blackhole	Presto の blackhole.properties ファイルの値を変更します。
presto-connector-cassandra	Presto の cassandra.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
presto-connector-jmx	Presto の jmx.properties ファイルの値を変更します。
presto-connector-kafka	Presto の kafka.properties ファイルの値を変更します。

分類	説明
presto-connector-localfile	Presto の localfile.properties ファイルの値を変更します。
presto-connector-mongodb	Presto の mongodb.properties ファイルの値を変更します。
presto-connector-mysql	Presto の mysql.properties ファイルの値を変更します。
presto-connector-postgresql	Presto の postgresql.properties ファイルの値を変更します。
presto-connector-raptor	Presto の raptor.properties ファイルの値を変更します。
presto-connector-redis	Presto の redis.properties ファイルの値を変更します。
presto-connector-tpch	Presto の tpch.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.7.4

### 4.7.4 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.148	0.147
Spark	1.6.2	1.6.2	1.6.2	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

#### 4.7.4 リリースノート

これは、Amazon S3 AWS へのリクエストの署名バージョン 4 認証を追加するためのパッチリリースです。アプリケーションとコンポーネントはすべて、以前の Amazon EMR リリースと同じです。

##### Important

このリリースバージョンでは、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。詳細については、「[新機能](#)」を参照してください。

#### 4.7.4 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	3.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.8.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。



コンポーネント	Version	説明
hadoop-client	2.7.2-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.2.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.1	HBase コマンドラインクライアント。
hbase-rest-server	1.2.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-6	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-6	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-6	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-6	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	1.0.0-amzn-6	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-6	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.46	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.148	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.148	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.2	Spark コマンドラインクライアント。
spark-history-server	1.6.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.3	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

#### 4.7.4 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.7.4 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-props-hive-site	HCatalog の <code>props-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の <code>log4j.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の <code>hive-exec-log4j.properties</code> ファイルの値を変更します。
hive-log4j	Hive の <code>hive-log4j.properties</code> ファイルの値を変更します。

分類	説明
hive-site	Hive の hive-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。



分類	説明
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.7.3

### 4.7.3 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.75
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.0	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.147	0.147
Spark	1.6.2	1.6.2	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

### 4.7.3 リリースノート

#### 4.7.3 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	3.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.8.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.2-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-3	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.7.2-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httfs-server	2.7.2-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.1	HBase コマンドラインクライアント。

コンポーネント	Version	説明
hbase-rest-server	1.2.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-6	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-6	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-6	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-6	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-6	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-6	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション

コンポーネント	Version	説明
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.46	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.148	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.148	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.2	Spark コマンドラインクライアント。
spark-history-server	1.6.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。



コンポーネント	Version	説明
spark-on-yarn	1.6.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.3	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

### 4.7.3 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## emr-4.7.3 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。

分類	説明
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。

分類	説明
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。

分類	説明
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。

分類	説明
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.7.2

### 4.7.2 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
AWS SDK for Java	1.10.75	1.10.75	1.10.75	1.10.27

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.0	0.12.0	0.11.1

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.147	0.147	0.143
Spark	1.6.2	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

## 4.7.2 リリースノート

以下のリリースノートには、Amazon EMR 4.7.2 に関する情報が含まれています。



リリース日: 2016 年 7 月 15 日

## 機能

- Mahout 0.12.2 にアップグレードしました
- Presto 0.148 にアップグレードしました
- Spark 1.6.2 にアップグレードしました
- パラメータとして URI を使用して、EMRFS AWSCredentialsProvider で使用する を作成できるようになりました。詳細については、[「EMRFS 用の を作成する AWSCredentialsProvider」](#) を参照してください。
- EMRFS では、ユーザーが `emrfs-site.xml` で `fs.s3.consistent.dynamodb.endpoint` プロパティを使用して、整合性のあるビューのメタデータのカスタム DynamoDB エンドポイントを設定できるようになりました。
- `/usr/bin` に `spark-example` というスクリプトを追加しました。これにより `/usr/lib/spark/spark/bin/run-example` をラップし、例を直接実行できます。例えば、Spark ディストリビューションに付属する SparkPi 例を実行するには、コマンドライン `spark-example SparkPi 100` から、または API のステップ `command-runner.jar` として を使用できます。

## 以前のリリースから解決された既知の問題

- Oozie で、Spark もインストールされたときに `spark-assembly.jar` が正しい場所がないために、Oozie で Spark アプリケーションを起動できなくなる問題を解決しました。
- YARN コンテナで Spark Log4j ベースのログインに関する問題を修正しました。

## 4.7.2 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオー

プンスソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	3.2.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.1.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.8.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	2.7.2-amzn-3	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-3	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-3	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-3	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-3	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-3	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-3	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-3	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-3	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。

コンポーネント	Version	説明
hbase-hmaster	1.2.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.1	HBase コマンドラインクライアント。
hbase-rest-server	1.2.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-6	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-6	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-6	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-6	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	1.0.0-amzn-6	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-6	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.2	機械学習用のライブラリ。
mysql-server	5.5.46	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.148	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.148	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.2	Spark コマンドラインクライアント。
spark-history-server	1.6.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.2	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.3	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.23	Apache HTTP サーバー。
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

## 4.7.2 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-4.7.2 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hadoop-ssl-server	Hadoop ssl のサーバー設定を変更します。
hadoop-ssl-client	Hadoop ssl のクライアント設定を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。

分類	説明
hbase-metrics	HBase の <code>hadoop-metrics2-hbase.properties</code> ファイルの値を変更します。
hbase-policy	HBase の <code>hbase-policy.xml</code> ファイルの値を変更します。
hbase-site	HBase の <code>hbase-site.xml</code> ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の <code>hdfs-site.xml</code> の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の <code>jndi.properties</code> の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の <code>proto-hive-site.xml</code> の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の <code>log4j.properties</code> の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の <code>webhcat-site.xml</code> ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の <code>hive-exec-log4j.properties</code> ファイルの値を変更します。
hive-log4j	Hive の <code>hive-log4j.properties</code> ファイルの値を変更します。



分類	説明
hive-site	Hive の hive-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。

分類	説明
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。

分類	説明
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.7.1

### 4.7.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
AWS SDK for Java	1.10.75	1.10.75	1.10.27	1.10.27
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.0	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.12.0	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.147	0.143	0.143
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	-	-
Trino (PrestoSQL)	-	-	-	-

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

### 4.7.1 リリースノート

以下のリリースノートには、Amazon EMR 4.7.1 に関する情報が含まれています。

リリース日: 2016 年 6 月 10 日

以前のリリースから解決された既知の問題

- プライベートサブネットを持つ VPC で起動されたクラスターの起動時間が長くなる問題を修正しました。このバグの影響があったのは、Amazon EMR 4.7.0 リリースで起動されたクラスターのみです。
- Amazon EMR 4.7.0 リリースで起動されたクラスターに対する Amazon EMR でのファイルのリスト処理が不適切であった問題を修正しました。

### 4.7.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオー

プンスソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	3.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.7.1	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	2.7.2-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hadoop-yarn-timeline-server	2.7.2-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。



コンポーネント	Version	説明
hbase-hmaster	1.2.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.1	HBase コマンドラインクライアント。
hbase-rest-server	1.2.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-5	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-5	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-5	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-5	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	1.0.0-amzn-5	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-5	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.0	機械学習用のライブラリ。
mysql-server	5.5.46	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.147	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。

コンポーネント	Version	説明
presto-worker	0.147	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.1	Spark コマンドラインクライアント。
spark-history-server	1.6.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.3	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.18	Apache HTTP サーバー。
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

#### 4.7.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.7.1 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。

分類	説明
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.7.0

### 4.7.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Phoenix](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Tez](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)



- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
AWS SDK for Java	1.10.75	1.10.27	1.10.27	1.10.27
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.0	1.2.0	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.143	0.143	0.140
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	-

## 4.7.0 リリースノート

### Important

Amazon EMR 4.7.0 は廃止されています。代わりに Amazon EMR 4.7.1 以降を使用してください。

リリース日: 2016 年 6 月 2 日

### 機能

- Apache Phoenix 4.7.0 を追加しました
- Apache Tez 0.8.3 を追加しました
- HBase 1.2.1 にアップグレードしました
- Mahout 0.12.0 にアップグレードしました
- Presto 0.147 にアップグレードしました
- AWS SDK for Java を 1.10.75 にアップグレードしました
- ユーザーがローカルモードで Pig を実行できるようにするため、最終的なフラグが `mapreduce.cluster.local.dir` の `mapred-site.xml` プロパティから削除されました。
- Amazon Redshift JDBC ドライバーがクラスターで利用可能

Amazon Redshift JDBC ドライバーが `/usr/share/aws/redshift/jdbc` に含まれました。/  
`usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar` は JDBC 4.1 互換の Amazon Redshift ドライバー、/  
`usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar` は JDBC 4.0 互換の Amazon Redshift ドライバーです。詳細については、「Amazon Redshift 管理ガイド」の  
「[JDBC 接続の設定](#)」を参照してください。

- Java 8

Presto を除き、OpenJDK 1.7 はすべてのアプリケーションに使用されるデフォルトの JDK です。ただし、OpenJDK 1.7 と 1.8 の両方がインストールされています。アプリケーションの

JAVA\_HOME を設定する方法については、「[Java 8 を使用したアプリケーションの設定](#)」を参照してください。

#### 以前のリリースから解決された既知の問題

- emr-4.6.0 で Amazon EMR 用のスループット最適化 HDD (st1) EBS ボリュームで著しくパフォーマンスに影響を与えていたカーネルの問題を修正しました。
- アプリケーションとして Hadoop を選択せずに HDFS 暗号化ゾーンを指定した場合にクラスターが失敗する問題を修正しました。
- デフォルトの HDFS 書き込みポリシーを RoundRobin から AvailableSpaceVolumeChoosingPolicy に変更しました。一部のボリュームは RoundRobin 設定で適切に使用されなかったため、コアノードに障害が発生し、HDFS の信頼性が低下しました。
- 整合性のあるビューを実現するためにデフォルトの DynamoDB メタデータテーブルを作成するときに、例外が発生する原因となっていた EMRFS CLI に関する問題を修正しました。
- マルチパートの名前の変更およびコピーオペレーション中に発生する可能性のあった、EMRFS のデッドロックの問題を修正しました。
- EMRFS で CopyPart サイズがデフォルト 5 MB になる問題を修正しました。現在では、デフォルト値は 128 MB で適切に設定されます。
- サービスを停止できなくなる可能性のある、Zeppelin upstart 設定の問題を修正しました。
- s3a:// がそれぞれのクラスパスで適切にロードされていないために、/usr/lib/hadoop/hadoop-aws.jar URI スキームを使用できなくなる Spark および Zeppelin の問題を修正しました。
- [HUE-2484](#) を移植しました。
- HBase ブラウザサンプルでの問題を修正するため、Hue 3.9.0 (JIRA が存在しない) から [commit](#) を移植しました。
- [HIVE-9073](#) を移植しました。

#### 4.7.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリ

ケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	3.1.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.2.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.4.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.7.1	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。

コンポーネント	Version	説明
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.2-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。

コンポーネント	Version	説明
hadoop-yarn-timeline-server	2.7.2-amzn-2	YARN アプリケーションの現在の情報と履歴情報を取得するためのサービス。
hbase-hmaster	1.2.1	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.1	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.1	HBase コマンドラインクライアント。
hbase-rest-server	1.2.1	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.1	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-5	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-5	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-5	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-5	Hive コマンドラインクライアント。

コンポーネント	Version	説明
hive-metastore-server	1.0.0-amzn-5	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-5	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-7	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.12.0	機械学習用のライブラリ。
mysql-server	5.5.46	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
phoenix-library	4.7.0-HBase-1.2	サーバーとクライアントの Phoenix ライブラリ
phoenix-query-server	4.7.0-HBase-1.2	JDBC のアクセス、プロトコルのバッファ、および Avatica API への JSON 形式のアクセスを提供するし軽量サーバー
presto-coordinator	0.147	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。



コンポーネント	Version	説明
presto-worker	0.147	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.1	Spark コマンドラインクライアント。
spark-history-server	1.6.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
tez-on-yarn	0.8.3	tez YARN アプリケーションおよびライブラリ。
webserver	2.4.18	Apache HTTP サーバー。
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

#### 4.7.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.7.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
phoenix-hbase-metrics	Phoenix の hadoop-metrics2-hbase.properties ファイルの値を変更します。
phoenix-hbase-site	Phoenix の hbase-site.xml ファイルの値を変更します。
phoenix-log4j	Phoenix の log4j.properties ファイルの値を変更します。

分類	説明
phoenix-metrics	Phoenix の hadoop-metrics2-phoenix.properties ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。

分類	説明
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
tez-site	Tez の tez-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.6.1

### 4.6.1 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)

- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
AWS SDK for Java	1.10.27	1.10.27	1.10.27	1.10.27
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	1.2.0	-	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.143	0.140	0.136
Spark	1.6.1	1.6.1	1.6.1	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-



	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
ZooKeeper-Sandbox	3.4.8	3.4.8	-	-

## 4.6.1 リリースノート

### 4.6.1 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	3.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクター。
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.1.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクター。
emr-s3-dist-cp	2.3.0	Amazon S3 に最適化された分散コピーアプリケーション。

コンポーネント	Version	説明
emrfs	2.6.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.2-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-1	HDFS オペレーションの HTTP エンドポイント。

コンポーネント	Version	説明
hadoop-kms-server	2.7.2-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hbase-hmaster	1.2.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.0	HBase コマンドラインクライアント。
hbase-rest-server	1.2.0	HBase の RESTful HTTP エンドポイントを提供するサービス。
hbase-thrift-server	1.2.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-4	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	1.0.0-amzn-4	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-4	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-4	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-4	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-4	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-6	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.11.1	機械学習用のライブラリ。
mysql-server	5.5	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
presto-coordinator	0.143	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.143	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.1	Spark コマンドラインクライアント。
spark-history-server	1.6.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
webserver	2.4	Apache HTTP サーバー。
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。

コンポーネント	Version	説明
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

#### 4.6.1 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.6.1 の分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hbase-env	HBase の環境の値を変更します。
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。

分類	説明
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。

分類	説明
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。



分類	説明
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。

分類	説明
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.6.0

### 4.6.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HBase](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Zeppelin-Sandbox](#)、[ZooKeeper-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
AWS SDK for Java	1.10.27	1.10.27	1.10.27	1.10.27
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	-	-	-
HCatalog	1.0.0	1.0.0	1.0.0	-
Hadoop	2.7.2	2.7.2	2.7.1	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.1	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.140	0.136	0.130
Spark	1.6.1	1.6.1	1.6.0	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	-	-	-

## 4.6.0 リリースノート

次のリリースノートには、Amazon EMR 4.6.0 リリースに関する情報が含まれています。

- HBase 1.2.0 を追加しました
- Zookeeper-Sandbox 3.4.8 を追加しました
- Presto-Sandbox 0.143 にアップグレードしました

- 現在、Amazon EMR リリースは Amazon Linux 2016.03.0 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/>」を参照してください。
- スループット最適化 HDD (st1) EBS ボリュームタイプに影響を及ぼす問題

Linux カーネルバージョン 4.2 以降の問題は、EMR 用のスループット最適化 HDD (st1) EBS ボリュームのパフォーマンスに大きな影響を及ぼします。このリリース (emr-4.6.0) ではカーネルバージョン 4.4.5 を使用するため、影響を受けます。したがって、st1 EBS ボリュームを使用する場合、emr-4.6.0 を使用しないことをお勧めします。emr-4.5.0 以前の Amazon EMR リリースと st1 であれば、影響を受けずに使用できます。これに加えて、将来のリリースで修正が提供されます。

- Python のデフォルト値

現在、Python 3.4 がデフォルトでインストールされますが、Python 2.7 はシステムデフォルトのままです。いずれかのブートストラップアクションを使用して、Python 3.4 をシステムのデフォルトとして設定できます。設定 API を使用して、が使用する Python バージョンに影響を与えるように spark-env、分類/usr/bin/python3.4で PYSPARK\_PYTHON エクスポートを に設定することができます PySpark。

- Java 8

Presto を除き、OpenJDK 1.7 はすべてのアプリケーションに使用されるデフォルトの JDK です。ただし、OpenJDK 1.7 と 1.8 の両方がインストールされています。アプリケーションの JAVA\_HOME を設定する方法については、「[Java 8 を使用したアプリケーションの設定](#)」を参照してください。

#### 以前のリリースから解決された既知の問題

- アプリケーションのプロビジョニングが、生成されたパスワードが原因でランダムに失敗する問題を修正しました。
- 以前は、mysqld がすべてのノードにインストールされました。現在では、選択されたアプリケーションにコンポーネントとして mysql-server が含まれている場合のみ、マスターインスタンスのみにインストールされます。現在、HCatalog、Hive、Hue、Presto-Sandbox、および Sqoop-Sandbox の各アプリケーションに、mysql-server コンポーネントが含まれています。
- yarn.scheduler.maximum-allocation-vcores をデフォルトの 32 から 80 に変更しました。これにより、コアインスタンスタイプが、YARN vcores が 32 より高く設定されているいくつかのラージインスタンスタイプのいずれかであるクラスターで、maximizeResourceAllocation オプションを使用中に Spark で主に

発生する、emr-4.4.0 での問題が修正されました。この問題の影響を受けていたのは、c4.8xlarge、cc2.8xlarge、hs1.8xlarge、i2.8xlarge、m2.4xlarge、r3.8xlarge、d2.8xlarge、または m4.10xlarge です。

- 現在では s3-dist-cp はすべての Amazon S3 候補に EMRFS を使用し、一時 HDFS ディレクトリにはステージングしません。
- クライアント側の暗号化のマルチパートアップロードの例外処理に関する問題を修正しました。
- ユーザーが Amazon S3 ストレージクラスを変更できるようにするオプションを追加しました。デフォルトでは、この設定は STANDARD です。emrfs-site 設定の分類設定は fs.s3.storageClass で、指定できる値は STANDARD、STANDARD\_IA、REDUCED\_REDUNDANCY です。ストレージクラスの詳細については、「Amazon Simple Storage Service ユーザーガイド」の「[ストレージクラス](#)」を参照してください。

#### 4.6.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、emr または aws で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion-amzn-EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の myapp-component というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは 2.2-amzn-2 として表示されます。

コンポーネント	Version	説明
emr-ddb	3.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.1.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.3.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.6.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.2-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-1	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.7.2-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-ftpfs-server	2.7.2-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hbase-hmaster	1.2.0	リージョンの調整および管理コマンドの実行を行う HBase クラスター用のサービス。
hbase-region-server	1.2.0	1 つ以上の HBase リージョンに対応するサービス。
hbase-client	1.2.0	HBase コマンドラインクライアント。
hbase-rest-server	1.2.0	HBase の RESTful HTTP エンドポイントを提供するサービス。



コンポーネント	Version	説明
hbase-thrift-server	1.2.0	HBase に Thrift エンドポイントを提供するサービス。
hcatalog-client	1.0.0-amzn-4	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。
hcatalog-server	1.0.0-amzn-4	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-4	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-4	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-4	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-4	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-6	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.11.1	機械学習用のライブラリ。

コンポーネント	Version	説明
mysql-server	5.5	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
presto-coordinator	0.143	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.143	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.1	Spark コマンドラインクライアント。
spark-history-server	1.6.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
webserver	2.4	Apache HTTP サーバー。

コンポーネント	Version	説明
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。
zookeeper-server	3.4.8	設定情報を維持し、名前を付け、分散化された同期を提供し、グループサービスを提供する一元化されたサービス。
zookeeper-client	3.4.8	ZooKeeper コマンドラインクライアント。

#### 4.6.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.6.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hbase-env	HBase の環境の値を変更します。

分類	説明
hbase-log4j	HBase の hbase-log4j.properties ファイルの値を変更します。
hbase-metrics	HBase の hadoop-metrics2-hbase.properties ファイルの値を変更します。
hbase-policy	HBase の hbase-policy.xml ファイルの値を変更します。
hbase-site	HBase の hbase-site.xml ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-props-hive-site	HCatalog の props-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。

分類	説明
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。

分類	説明
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。
sqoop-oraoop-site	Sqoop OraOop の oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。

分類	説明
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。
zookeeper-config	ZooKeeperの zoo.cfg ファイルの値を変更します。
zookeeper-log4j	ZooKeeperの log4j.properties ファイルの値を変更します。

## Amazon EMR リリース 4.5.0

### 4.5.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Zeppelin-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の3つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
AWS SDK for Java	1.10.27	1.10.27	1.10.27	1.10.27
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.6.0
HBase	-	-	-	-
HCatalog	1.0.0	1.0.0	-	-
Hadoop	2.7.2	2.7.1	2.7.1	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-



	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.140	0.136	0.130	0.125
Spark	1.6.1	1.6.0	1.6.0	1.5.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.5	0.5.5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

## 4.5.0 リリースノート

次のリリースノートには、Amazon EMR 4.5.0 リリースに関する情報が含まれています。

リリース日: 2016 年 4 月 4 日

### 機能

- Spark 1.6.1 にアップグレードしました
- Hadoop 2.7.2 にアップグレードしました
- Presto 0.140 にアップグレードしました
- Amazon S3 サーバー側の暗号化 AWS KMS のサポートが追加されました。

### 以前のリリースから解決された既知の問題

- ノードが再起動された後に MySQL および Apache サーバーが起動しない問題を修正しました。
- Amazon S3 に保存されているパーティション分割されていないテーブルで IMPORT が正しく機能しない問題を修正しました。
- Hive テーブルに書き込むときに、ステージングディレクトリが /mnt/tmp ではなく /tmp であることが要求される Presto の問題を修正しました。

## 4.5.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	3.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.1.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.2.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.5.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	2.7.2-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.2-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.2-amzn-0	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.2-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.2-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.2-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.2-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.2-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.2-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hcatalog-client	1.0.0-amzn-4	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。

コンポーネント	Version	説明
hcatalog-server	1.0.0-amzn-4	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-4	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-4	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-4	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-4	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-5	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.11.1	機械学習用のライブラリ。
mysql-server	5.5	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
presto-coordinator	0.140	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.140	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.1	Spark コマンドラインクライアント。
spark-history-server	1.6.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.1	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
webserver	2.4	Apache HTTP サーバー。
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

## 4.5.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-4.5.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。

分類	説明
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。



分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

## Amazon EMR リリース 4.4.0

### 4.4.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[HCatalog](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Sqoop-Sandbox](#)、[Zeppelin-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
AWS SDK for Java	1.10.27	1.10.27	1.10.27	追跡されていません
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.6.0	-
HBase	-	-	-	-
HCatalog	1.0.0	-	-	-
Hadoop	2.7.1	2.7.1	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
MXNet	-	-	-	-
Mahout	0.11.1	0.11.0	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.0.1
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.136	0.130	0.125	0.119
Spark	1.6.0	1.6.0	1.5.2	1.5.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.5	0.5.5	0.6.0-SNA PSHOT
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

## 4.4.0 リリースノート

次のリリースノートには、Amazon EMR 4.4.0 リリースに関する情報が含まれています。

リリース日: 2016 年 3 月 14 日

### 機能

- HCatalog 1.0.0 を追加しました
- Sqoop-Sandbox 1.4.6 を追加しました
- Presto 0.136 にアップグレードしました
- Zeppelin 0.5.6 にアップグレードしました
- Mahout 0.11.1 にアップグレードしました
- デフォルトで `dynamicResourceAllocation` を有効にしました。
- リリースのすべての設定分類の表を追加しました。詳細については、「[アプリケーションの設定](#)」の設定分類の表を参照してください。

### 以前のリリースから解決された既知の問題

- `maximizeResourceAllocation` 設定で YARN ApplicationMaster デーモン用に十分なメモリが予約されない問題を修正しました。
- カスタム DNS で発生した問題を修正しました。`resolve.conf` のエントリが、提供されたカスタムエントリよりも前に指定されている場合、そのカスタムエントリは解決されません。この動作は、デフォルトの VPC ネームサーバーが `resolve.conf` のトップエントリとして挿入された VPC のクラスターによって影響を受けました。
- デフォルトの Python がバージョン 2.7 に移行した場合に、そのバージョンに対して boto がインストールされなかった問題を修正しました。
- YARN コンテナと Spark アプリケーションが独自の Ganglia ラウンドロビンデータベース (rrd) ファイルを生成し、それによりインスタンスにアタッチされた最初のディスクがいっぱいになる問題を修正しました。この修正によって、YARN コンテナレベルのメトリクスが無効になり、Spark アプリケーションレベルのメトリクスが無効になりました。
- ログプッシャーですべての空のログフォルダーが削除される問題を修正しました。この問題により、ログプッシャーは user の空の `/var/log/hive` フォルダを削除したため、Hive CLI はログを記録できませんでした。
- パーティション分割に影響し、インポート中にエラーを発生させた、Hive のインポートに影響を与える問題を修正しました。

- EMRFS と s3-dist-cp が、ピリオドを含むバケット名を適切に処理しなかった問題を修正しました。
- EMRFS の動作を変更し、バージョニングが有効なバケットで、`_${folder}$` マーカーファイルが連続して作成されないようにしました。これにより、バージョニングが有効なバケットでパフォーマンスが向上する可能性があります。
- クライアント側の暗号化が有効になっている場合を除き、インストラクションファイルを使用しないよう EMRFS の動作を変更しました。クライアント側の暗号化を使用中にインストラクションファイルを削除する場合は、`emrfs-site.xml` プロパティの `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled` を `true` に設定できます。
- YARN ログの集計を変更し、集計先でログを 2 日間保持するようにしました。デフォルトの送信先はクラスターの HDFS ストレージです。この期間を変更する場合は、クラスターの作成時に `yarn.log-aggregation.retain-seconds` 設定分類を使用して `yarn-site` の値を変更します。これまでどおり、クラスターの作成時は、`log-uri` パラメータを使用して、Amazon S3 にアプリケーションログを保存できます。

#### 適用されたパッチ

- [HIVE-9655](#)
- [HIVE-9183](#)
- [HADOOP-12810](#)

#### 4.4.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	3.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.1.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.2.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.4.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。

コンポーネント	Version	説明
hadoop-client	2.7.1-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.1-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.1-amzn-1	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.7.1-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.7.1-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.1-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.1-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.1-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.1-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hcatalog-client	1.0.0-amzn-3	hcatalog-server を操作するための 'hcat' コマンドラインクライアント。



コンポーネント	Version	説明
hcatalog-server	1.0.0-amzn-3	分散アプリケーション用のテーブルおよびストレージ管理レイヤーである HCatalog を提供するサービス。
hcatalog-webhcat-server	1.0.0-amzn-3	HCatalog に REST インターフェイスを提供する HTTP エンドポイント。
hive-client	1.0.0-amzn-3	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-3	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-3	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-5	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.11.1	機械学習用のライブラリ。
mysql-server	5.5	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。

コンポーネント	Version	説明
presto-coordinator	0.136	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.136	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.0	Spark コマンドラインクライアント。
spark-history-server	1.6.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.0	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.6.0	YARN スレーブで必要な Apache Spark ライブラリ。
sqoop-client	1.4.6	Apache Sqoop コマンドラインクライアント。
webserver	2.4	Apache HTTP サーバー。
zeppelin-server	0.5.6-incubating	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

## 4.4.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-4.4.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hcatalog-env	HCatalog 環境の値を変更します。
hcatalog-server-jndi	HCatalog の jndi.properties の値を変更します。
hcatalog-server-proto-hive-site	HCatalog の proto-hive-site.xml の値を変更します。
hcatalog-webhcat-env	HCatalog WebHCat 環境の値を変更します。
hcatalog-webhcat-log4j	HCatalog WebHCat の log4j.properties の値を変更します。

分類	説明
hcatalog-webhcat-site	HCatalog WebHCat の webhcat-site.xml ファイルの値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。

分類	説明
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
sqoop-env	Sqoop の環境の値を変更します。

分類	説明
sqoop-oraoop-site	Sqoop OraOopの oraoop-site.xml ファイルの値を変更します。
sqoop-site	Sqoop の sqoop-site.xml ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

## Amazon EMR リリース 4.3.0

### 4.3.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Zeppelin-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

## アプリケーションバージョン情報

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK for Java	1.10.27	1.10.27	追跡されていません	追跡されていません
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNA PSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-



### 4.3.0 リリースノート

次のリリースノートには、Amazon EMR 4.3.0 リリースに関する情報が含まれています。

リリース日: 2016 年 1 月 19 日

#### 機能

- Hadoop 2.7.1 にアップグレードしました
- Spark 1.6.0 にアップグレードしました
- Ganglia を 3.7.2 にアップグレードしました
- Presto を 0.130 にアップグレードしました
- Amazon EMR を true に設定すると、`spark.dynamicAllocation.enabled` にくつかわりの変更が加えられていました。デフォルトでは false です。true に設定すると、`maximizeResourceAllocation` 設定で定義されているデフォルト設定に影響を与えます。
- `spark.dynamicAllocation.enabled` を true に設定した場合、`spark.executor.instances` は `maximizeResourceAllocation` によって設定されません。
- `spark.driver.memory` 設定は、`spark.executors.memory` 設定と同様に、クラスター内のインスタンスタイプに基づいて定義されます。ただし、Spark ドライバーアプリケーションは、マスターインスタンスまたはいずれかのコアインスタンスで (たとえば、YARN クライアントモードとクラスターモードのそれぞれで) 実行されるため、`spark.driver.memory` 設定は、これらの 2 つのインスタンスグループ間で、小さい方のインスタンスのインスタンスタイプに基づいて定義されます。
- `spark.default.parallelism` 設定は、YARN コンテナに使用可能な CPU コアの数の 2 倍に定義されます。以前のリリースでは、半分の値に定義されていました。
- Spark YARN プロセス用に予約されるメモリアーヘッドの計算精度が上がったため、Spark に使用可能なメモリの合計量 (`spark.executor.memory`) がわずかに増えました。

#### 以前のリリースから解決された既知の問題

- 現在、YARN ログの集計はデフォルトで有効になります。
- YARN ログの集計が有効な場合に、クラスターの Amazon S3 ログバケットにログがプッシュされない問題を修正しました。
- YARN コンテナサイズは、すべてのノードタイプで新たに最低 32 になりました。

- 大規模なクラスターのプライマリノードで過剰なディスク I/O を発生させる Ganglia の問題を修正しました。
- クラスターのシャットダウン時に Amazon S3 にアプリケーションログがプッシュされない問題を修正しました。
- 特定のコマンドを失敗させる EMRFS CLI の問題を修正しました。
- Zeppelin で、基盤となる に依存関係がロードされない問題を修正しました SparkContext。
- インスタンスの追加を試みるサイズ変更の発行によって発生する問題を修正しました。
- CREATE TABLE AS SELECT が Amazon S3 への過剰なリスト呼び出しを行う Hive の問題を修正しました。
- Hue、Oozie、および Ganglia がインストールされていると、大規模なクラスターが適切にプロビジョニングされない問題を修正しました。
- エラーで失敗した場合でもゼロ終了コードを返す s3-dist-cp の問題を修正しました。

#### 4.3.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	3.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。

コンポーネント	Version	説明
emr-kinesis	3.1.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.1.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.3.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.7.2	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.7.2	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.7.1	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.7.1-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.7.1-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-library	2.7.1-amzn-0	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.7.1-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-https-server	2.7.1-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.7.1-amzn-0	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.7.1-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.7.1-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.7.1-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hive-client	1.0.0-amzn-2	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-2	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-2	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	3.7.1-amzn-5	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.11.0	機械学習用のライブラリ。
mysql-server	5.5	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
presto-coordinator	0.130	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.130	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.6.0	Spark コマンドラインクライアント。
spark-history-server	1.6.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.6.0	YARN のメモリ内実行エンジン。

コンポーネント	Version	説明
spark-yarn-slave	1.6.0	YARN スレーブに必要な Apache Spark ライブラリ。
webserver	2.4	Apache HTTP サーバー。
zeppelin-server	0.5.5-incubating-amzn-1	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

### 4.3.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

#### emr-4.3.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。

分類	説明
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。



## Amazon EMR リリース 4.2.0

### 4.2.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Ganglia](#)、[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Zeppelin-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK for Java	1.10.27	1.10.27	追跡されていません	追跡されていません
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNA PSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

## 4.2.0 リリースノート

次のリリースノートには、Amazon EMR 4.2.0 リリースに関する情報が含まれています。

リリース日: 2015 年 11 月 18 日

### 機能

- Ganglia のサポートを追加しました
- Spark 1.5.2 にアップグレードしました
- Presto 0.125 にアップグレードしました
- Oozie を 4.2.0 にアップグレードしました
- Zeppelin を 0.5.5 にアップグレードしました
- を 1.10.27 AWS SDK for Java にアップグレードしました

## 以前のリリースから解決された既知の問題

- デフォルトのメタデータテーブル名を使用しない EMRFS CLI の問題を修正しました。
- Amazon S3 で ORC-backed テーブルを使用するときに発生した問題を修正しました。
- Spark 設定で Python バージョンが一致しない問題を修正しました。
- VPC のクラスターでの DNS の問題により、YARN ノードのステータスが報告されない問題を修正しました。
- YARN がノードを廃棄することが原因でアプリケーションがハングしたり、新しいアプリケーションを予定できなくなったりする問題を修正しました。
- クラスターが TIMED\_OUT\_STARTING というステータスで終了するときに発生する問題を修正しました。
- EMRFS Scala 依存関係を他のビルドに含める場合に発生する問題を修正しました。Scala 依存関係が削除されました。

## 4.2.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	3.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。

コンポーネント	Version	説明
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.1.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.0.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.2.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
ganglia-monitor	3.6.0	Ganglia モニタリングエージェントとともに埋め込まれた、Hadoop エコシステムアプリケーション用 Ganglia エージェント。
ganglia-metadata-collector	3.6.0	Ganglia モニタリングエージェントからメトリクスを集計する Ganglia メタデータコレクター。
ganglia-web	3.5.10	Ganglia メタデータコレクターによって収集されたメトリクスを表示するウェブアプリケーション。
hadoop-client	2.6.0-amzn-2	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.6.0-amzn-2	ブロックを保存する HDFS ノードレベルのサービス。

コンポーネント	Version	説明
hadoop-hdfs-library	2.6.0-amzn-2	HDFS コマンドラインクライアントとライブラリ
hadoop-hdfs-namenode	2.6.0-amzn-2	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httfs-server	2.6.0-amzn-2	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.6.0-amzn-2	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.6.0-amzn-2	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.6.0-amzn-2	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.6.0-amzn-2	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hive-client	1.0.0-amzn-1	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。

コンポーネント	Version	説明
hive-server	1.0.0-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。
hue-server	3.7.1-amzn-5	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.11.0	機械学習用のライブラリ。
mysql-server	5.5	MySQL データベースサーバー。
oozie-client	4.2.0	Oozie コマンドラインクライアント。
oozie-server	4.2.0	Oozie ワークフローリクエストを受け入れるサービス。
presto-coordinator	0.125	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.125	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.5.2	Spark コマンドラインクライアント。
spark-history-server	1.5.2	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。

コンポーネント	Version	説明
spark-on-yarn	1.5.2	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.5.2	YARN スレーブで必要な Apache Spark ライブラリ。
webserver	2.4	Apache HTTP サーバー。
zeppelin-server	0.5.5-incubating-amzn-0	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

## 4.2.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

### emr-4.2.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。



分類	説明
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します。
hue-ini	Hue の ini ファイルの値を変更します。
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。

分類	説明
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
presto-connector-hive	Presto の hive.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
spark-metrics	Spark の metrics.properties ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。

分類	説明
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

## Amazon EMR リリース 4.1.0

### 4.1.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Hadoop](#)、[Hive](#)、[Hue](#)、[Mahout](#)、[Oozie-Sandbox](#)、[Pig](#)、[Presto-Sandbox](#)、[Spark](#)、[Zeppelin-Sandbox](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

### アプリケーションバージョン情報

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK for Java	1.10.27	1.10.27	追跡されていません	追跡されていません
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNA PSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

## 4.1.0 リリースノート

### 4.1.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通

常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、*CommunityVersion*-amzn-*EmrVersion* の形式のバージョンラベルがあります。*EmrVersion* は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
<code>emr-ddb</code>	3.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
<code>emr-goodies</code>	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
<code>emr-kinesis</code>	3.1.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
<code>emr-s3-dist-cp</code>	2.0.0	Amazon S3 に最適化された分散コピーアプリケーション。
<code>emrfs</code>	2.1.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
<code>hadoop-client</code>	2.6.0-amzn-1	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
<code>hadoop-hdfs-datanode</code>	2.6.0-amzn-1	ブロックを保存する HDFS ノードレベルのサービス。
<code>hadoop-hdfs-library</code>	2.6.0-amzn-1	HDFS コマンドラインクライアントとライブラリ

コンポーネント	Version	説明
hadoop-hdfs-namenode	2.6.0-amzn-1	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-ftpfs-server	2.6.0-amzn-1	HDFS オペレーションの HTTP エンドポイント。
hadoop-kms-server	2.6.0-amzn-1	Hadoop の KeyProvider API に基づく暗号化キー管理サーバー。
hadoop-mapred	2.6.0-amzn-1	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。
hadoop-yarn-nodemanager	2.6.0-amzn-1	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.6.0-amzn-1	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hive-client	1.0.0-amzn-1	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-1	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-1	Hive クエリをウェブリクエストとして受け入れるサービス。

コンポーネント	Version	説明
hue-server	3.7.1-amzn-4	Hadoop エコシステムアプリケーションを使用してデータを分析するウェブアプリケーション
mahout-client	0.11.0	機械学習用のライブラリ。
mysql-server	5.5	MySQL データベースサーバー。
oozie-client	4.0.1	Oozie コマンドラインクライアント。
oozie-server	4.0.1	Oozie ワークフローリクエストを受け入れるサービス。
presto-coordinator	0.119	presto-worker 間でクエリを受け入れ、クエリの実行を管理するサービス。
presto-worker	0.119	いくつかのクエリを実行するサービス。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.5.0	Spark コマンドラインクライアント。
spark-history-server	1.5.0	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。
spark-on-yarn	1.5.0	YARN のメモリ内実行エンジン。



コンポーネント	Version	説明
spark-yarn-slave	1.5.0	YARN スレーブに必要な Apache Spark ライブラリ。
zeppelin-server	0.6.0-incubating-SNAPSHOT	インタラクティブなデータ分析を可能にするウェブベースのノートブック。

#### 4.1.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.1.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hdfs-encryption-zones	HDFS 暗号化ゾーンを設定します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hive-env	Hive 環境の値を変更します。

分類	説明
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
hue-ini	Hue の ini ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
hadoop-kms-acls	Hadoop の kms-acls.xml ファイルの値を変更します。
hadoop-kms-env	Hadoop KMS 環境の値を変更します。
hadoop-kms-log4j	Hadoop の kms-log4j.properties ファイルの値を変更します。
hadoop-kms-site	Hadoop の kms-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
oozie-env	Oozie の環境の値を変更します。
oozie-log4j	Oozie の oozie-log4j.properties ファイルの値を変更します。

分類	説明
oozie-site	Oozie の oozie-site.xml ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
presto-log	Presto の log.properties ファイルの値を変更します。
presto-config	Presto の config.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。
zeppelin-env	Zeppelin 環境の値を変更します。

## Amazon EMR リリース 4.0.0

### 4.0.0 アプリケーションバージョン

このリリースでは、次のアプリケーションがサポートされています:

[Hadoop](#)、[Hive](#)、[Mahout](#)、[Pig](#)、[Spark](#)。

以下の表は、Amazon EMR のこのリリースで利用可能なアプリケーションバージョンと、前の 3 つの Amazon EMR リリース (該当する場合) で利用可能なアプリケーションバージョンを示しています。

Amazon EMR の各リリースのアプリケーションバージョンの包括的な履歴については、以下のトピックを参照してください。

- [Amazon EMR 7.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 6.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 5.x リリースのアプリケーションバージョン](#)
- [Amazon EMR 4.x リリースのアプリケーションバージョン](#)

#### アプリケーションバージョン情報

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK for Java	1.10.27	1.10.27	追跡されていません	追跡されていません
Python	追跡されていません	追跡されていません	追跡されていません	追跡されていません
Scala	追跡されていません	追跡されていません	追跡されていません	追跡されていません
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNA PSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

## 4.0.0 リリースノート

### 4.0.0 コンポーネントバージョン

このリリースで Amazon EMR がインストールするコンポーネントを以下に示します。そのうちいくつかは、ビッグデータアプリケーションパッケージの一部としてインストールされます。その他は Amazon EMR に固有であり、システムプロセスと機能に対してインストールされます。これらは通常、`emr` または `aws` で開始されます。通常、最新の Amazon EMR リリースのビッグデータアプリケーションパッケージは、コミュニティにある最新バージョンです。コミュニティリリースは、できるだけ早く Amazon EMR で入手可能になるようにしています。

Amazon EMR の一部のコンポーネントは、コミュニティバージョンとは異なります。これらのコンポーネントには、`CommunityVersion-amzn-EmrVersion` の形式のバージョンラベルがあります。`EmrVersion` は 0 から始まります。例えば、バージョン 2.2 の `myapp-component` というオープンソースコミュニティコンポーネントが、異なる Amazon EMR リリースに組み込むために 3 回変更された場合、そのリリースバージョンは `2.2-amzn-2` として表示されます。

コンポーネント	Version	説明
emr-ddb	3.0.0	Hadoop エコシステムアプリケーション用の Amazon DynamoDB コネクタ。
emr-goodies	2.0.0	Hadoop エコシステムに役立つ追加のライブラリ。
emr-kinesis	3.0.0	Hadoop エコシステムアプリケーション用の Amazon Kinesis コネクタ。
emr-s3-dist-cp	2.0.0	Amazon S3 に最適化された分散コピーアプリケーション。
emrfs	2.0.0	Hadoop エコシステムアプリケーション用の Amazon S3 コネクタ。
hadoop-client	2.6.0-amzn-0	'hdfs'、'hadoop'、'yarn' などの Hadoop コマンドラインクライアント。
hadoop-hdfs-datanode	2.6.0-amzn-0	ブロックを保存する HDFS ノードレベルのサービス。
hadoop-hdfs-namenode	2.6.0-amzn-0	ファイル名を追跡し、場所をブロックする HDFS サービス。
hadoop-httpfs-server	2.6.0-amzn-0	HDFS オペレーションの HTTP エンドポイント。
hadoop-mapred	2.6.0-amzn-0	MapReduce MapReduce アプリケーションを実行するための実行エンジンライブラリ。

コンポーネント	Version	説明
hadoop-yarn-nodemanager	2.6.0-amzn-0	個別のノードでコンテナを管理する YARN サービス。
hadoop-yarn-resourcemanager	2.6.0-amzn-0	クラスターリソースおよび分散アプリケーションを割り当て、管理する YARN サービス。
hive-client	1.0.0-amzn-0	Hive コマンドラインクライアント。
hive-metastore-server	1.0.0-amzn-0	Hadoop オペレーションの SQL 用のメタデータを保存するセマンティックレポジトリである Hive メタストアにアクセスするためのサービス。
hive-server	1.0.0-amzn-0	Hive クエリをウェブリクエストとして受け入れるサービス。
mahout-client	0.10.0	機械学習用のライブラリ。
mysql-server	5.5	MySQL データベースサーバー。
pig-client	0.14.0-amzn-0	Pig コマンドラインクライアント。
spark-client	1.4.1	Spark コマンドラインクライアント。
spark-history-server	1.4.1	完了した Spark アプリケーションの有効期間にログに記録されたイベントを表示するウェブ UI。



コンポーネント	Version	説明
spark-on-yarn	1.4.1	YARN のメモリ内実行エンジン。
spark-yarn-slave	1.4.1	YARN スレーブで必要な Apache Spark ライブラリ。

#### 4.0.0 設定分類

設定分類を使用すると、アプリケーションをカスタマイズできます。これらは多くの場合、hive-site.xml などのアプリケーションの構成 XML ファイルに対応します。詳細については、「[アプリケーションの設定](#)」を参照してください。

##### emr-4.0.0 分類

分類	説明
capacity-scheduler	Hadoop の capacity-scheduler.xml ファイルの値を変更します。
core-site	Hadoop の core-site.xml ファイルの値を変更します。
emrfs-site	EMRFS の設定を変更します。
hadoop-env	Hadoop のすべてのコンポーネントに対する Hadoop 環境の値を変更します。
hadoop-log4j	Hadoop の log4j.properties ファイルの値を変更します。
hdfs-site	HDFS の hdfs-site.xml の値を変更します。
hive-env	Hive 環境の値を変更します。
hive-exec-log4j	Hive の hive-exec-log4j.properties ファイルの値を変更します。

分類	説明
hive-log4j	Hive の hive-log4j.properties ファイルの値を変更します。
hive-site	Hive の hive-site.xml ファイルの値を変更します
httpfs-env	HTTPFS 環境の値を変更します。
httpfs-site	Hadoop の httpfs-site.xml ファイルの値を変更します。
mapred-env	MapReduce アプリケーションの環境の値を変更します。
mapred-site	MapReduce アプリケーションの mapred-site.xml ファイルの値を変更します。
pig-properties	Pig の pig.properties ファイルの値を変更します。
pig-log4j	Pig の log4j.properties ファイルの値を変更します。
spark	Apache Spark 用の Amazon EMR-curated 設定。
spark-defaults	Spark の spark-defaults.conf ファイルの値を変更します。
spark-env	Spark 環境の値を変更します。
spark-log4j	Spark の log4j.properties ファイルの値を変更します。
yarn-env	YARN 環境の値を変更します。
yarn-site	YARN の yarn-site.xml ファイルの値を変更します。

# Amazon EMR 2.x および 3.x AMI バージョン

## Note

AWS は、すべての AWS API エンドポイントの TLS 設定を TLS 1.2 の最小バージョンに更新しています。Amazon EMR リリース 3.10 以前では、TLS 1.0/1.1 接続のみがサポートされています。2023 年 12 月 4 日以降は、Amazon EMR 3.10 以前のバージョンでクラスターを作成できなくなります。

Amazon EMR 3.10 以前を使用している場合は、すぐにワークロードをテストし、最新の Amazon EMR リリースに移行することをお勧めします。詳細については、「[AWS Security Blog](#)」を参照してください。

Amazon EMR 2.x および 3.x リリースは、AMI バージョンと呼ばれ、互換性の理由からそれらを必要とする既存のソリューションで利用できます。これらのリリースバージョンで、新しいクラスターまたは新しいソリューションを作成することはお勧めしません。これらには、新しいリリースの機能がなく、古いアプリケーションパッケージが含まれています。

最新の Amazon EMR リリースバージョンを使用してソリューションを構築することをお勧めします。

2.x および 3.x シリーズのリリースバージョンと最新の Amazon EMR リリースバージョンの違いの範囲は重要です。これらの相違点は、クラスターの作成方法や設定方法から、クラスターのアプリケーションのポートとディレクトリの構造まで、多岐にわたります。

このセクションでは、Amazon EMR の最も重要な相違点と、特定のアプリケーションの設定と管理の違いについて説明します。包括的ではありません。2.x または 3.x シリーズでクラスターを作成して使用すると、このセクションで説明されていない相違点が発生することがあります。

## トピック

- [Amazon EMR の以前の AMI バージョンでクラスターを作成する](#)
- [Amazon EMR の以前の AMI のバージョンを使用したアプリケーションのインストール](#)
- [Amazon EMR の以前の AMI バージョンによるクラスターおよびアプリケーション設定のカスタマイズ](#)
- [Amazon EMR の以前の AMI バージョンの Hive アプリケーションの仕様。](#)
- [Amazon EMR の以前の AMI バージョンの HBase アプリケーションの仕様](#)
- [Amazon EMR の以前の AMI バージョンの Pig アプリケーションの仕様。](#)

- [Amazon EMR の以前の AMI バージョンの Spark アプリケーションの仕様。](#)
- [Amazon EMR の以前の AMI バージョンとの S3DistCp ユーティリティの違い](#)

## Amazon EMR の以前の AMI バージョンでクラスターを作成する

Amazon EMR 2.x および 3.x リリースは AMI バージョンによって参照されます。Amazon EMR リリース 4.0.0 以降では、リリースは `emr-5.11.0` などのリリースラベルを使用してリリースバージョンによって参照されます。この変更は、AWS CLI またはプログラムでクラスターを作成するときに最も顕著です。

を使用して AMI リリースバージョンを使用してクラスター AWS CLI を作成する場合は、などの `--ami-version` オプションを使用します `--ami-version 3.11.0`。Amazon EMR 4.0.0 以降で導入された多くのオプション、機能、およびアプリケーションは、`--ami-version` を指定すると利用できません。詳細については、「AWS CLI コマンドリファレンス」の「[create-cluster](#)」を参照してください。

次の AWS CLI コマンド例では、AMI バージョンを使用してクラスターを起動します。

### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \  
--applications Name=Hue Name=Hive Name=Pig \  
--use-default-roles --ec2-attributes KeyName=myKey \  
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,\  
InstanceType=m3.xlarge InstanceGroupType=CORE,InstanceCount=2,\  
InstanceType=m3.xlarge --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-  
actions/configure-hadoop,\  
Name="Configuring infinite JVM reuse",Args=["-m", "mapred.job.reuse.jvm.num.tasks=-1"]
```

プログラムでは、すべての Amazon EMR リリースバージョンが EMR API でクラスターを作成する `RunJobFlowRequest` アクションを使用します。次の Java コード例では、AMI リリースバージョン 3.11.0 を使用してクラスターを作成します。

```
RunJobFlowRequest request = new RunJobFlowRequest()
```

```
.withName("AmiVersion Cluster")
.withAmiVersion("3.11.0")
.withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

次の `RunJobFlowRequest` 呼び出しには、代わりにリリースラベルを使用します。

```
RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("ReleaseLabel Cluster")
    .withReleaseLabel("emr-7.1.0")
    .withInstances(new JobFlowInstancesConfig()
        .withEc2KeyName("myKeyPair")
        .withInstanceCount(1)
        .withKeepJobFlowAliveWhenNoSteps(true)
        .withMasterInstanceType("m3.xlarge")
        .withSlaveInstanceType("m3.xlarge");
```

## クラスターサイズの設定

クラスターの実行時、Hadoop は、データを処理するときに必要だったマッパータスクとリデューサータスクの数を確認します。クラスターが大きければ、リソースを効率的に使用し、処理時間を短縮するために、タスクの数は多くなるはずですが、通常、EMR クラスターのサイズは、そのクラスターを通して変わりません。タスクの数はクラスターの作成時に設定します。実行中のクラスターのサイズを変更すると、クラスター実行中に処理を変更できます。つまり、タスクの数は固定されておらず、クラスターが存続している間は変更できます。最適な数のタスクを設定するのに役立つ設定オプションは 2 つあります：

- `mapred.map.tasksperslot`
- `mapred.reduce.tasksperslot`

これらのオプションは両方とも `mapred-conf.xml` ファイルで設定できます。ジョブをクラスターに送信するとき、クラスター全体で使用できるマップおよびリデューススロットの現在の合計数がジョブクライアントによって確認されます。その後、ジョブクライアントは次の方程式を使用して、タスクの数を設定します。

- `mapred.map.tasks = mapred.map.tasksperslot × クラスターのマップスロット数`

- `mapred.reduce.tasks = mapred.reduce.tasksperslot × クラスターのリデューススロット数`

ジョブクライアントは、タスク数が設定されていない場合にのみ `tasksperslot` パラメータを読み込みます。タスク数はいつでも無効にできます。この無効にする処理は、ブートストラップアクションを使用してすべてのクラスターに対してまとめて実行することも、設定を変更するステップを追加してジョブごとに個別に実行することも可能です。

Amazon EMR は、タスクノードエラーに対する耐性があり、タスクノードが使用できなくなってもクラスターの実行を継続できます。Amazon EMR は、追加のタスクノードを自動的にプロビジョニングし、エラーが発生したタスクノードの代わりに配置します。

クラスターステップごとに異なる数のタスクノードを設定できます。また、実行中のクラスターにステップを追加して、タスクノードの数を変更することもできます。デフォルトで、すべてのステップが確実に順次実行されるので、どのステップに対しても、実行中のタスクノードの数を指定できます。

## Amazon EMR の以前の AMI のバージョンを使用したアプリケーションのインストール

AMI バージョンを使用する場合、アプリケーションは、[RunJobFlow](#) アクションの `NewSupportedProducts` パラメータの使用、ブートストラップアクションの使用、[ステップ](#) アクションの使用など、さまざまな方法でインストールされます。

## Amazon EMR の以前の AMI バージョンによるクラスターおよびアプリケーション設定のカスタマイズ

Amazon EMR リリースバージョン 4.0.0 では、設定分類を使用してアプリケーションを設定する簡単な方法が導入されました。詳細については、「[アプリケーションの設定](#)」を参照してください。AMI バージョンを使用する場合は、渡した引数とともにブートストラップアクションを使用してアプリケーションを設定します。たとえば、ブートストラップアクション `configure-hadoop` および `configure-daemons` は、Hadoop や YARN に固有の `--namenode-heap-size` などの環境プロパティを設定します。より新しいバージョンでは、これらは `hadoop-env` と `yarn-env` 設定分類を使用して設定されます。一般的な設定を実行するブートストラップアクションについては、[emr-bootstrap-actions Github のリポジトリ](#)を参照してください。

次の表は、より最近の Amazon EMR リリースバージョンでブートストラップアクションを構成分類にマップしています。

## Hadoop

影響を受けるアプリケーションファイル名	AMI バージョンのブートストラップアクション	設定分類
core-site.xml	configure-hadoop -c	core-site
log4j.properties	configure-hadoop -l	hadoop-log4j
hdfs-site.xml	configure-hadoop -s	hdfs-site
該当なし	該当なし	hdfs-encryption-zones
mapred-site.xml	configure-hadoop -m	mapred-site
yarn-site.xml	configure-hadoop -y	yarn-site
httpfs-site.xml	configure-hadoop -t	httpfs-site
capacity-scheduler.xml	configure-hadoop -z	capacity-scheduler
yarn-env.sh	configure-daemons --resource-manager-opts	yarn-env

## [Hive]

影響を受けるアプリケーションファイル名	AMI バージョンのブートストラップアクション	設定分類
hive-env.sh	該当なし	hive-env
hive-site.xml	hive-script --install -hive-site \${MY_HIVE_SITE_FILE}	hive-site
hive-exec-log4j.properties	該当なし	hive-exec-log4j

影響を受けるアプリケーションファイル名	AMI バージョンのブートストラップアクション	設定分類
hive-log4j.properties	該当なし	hive-log4j

## EMRFS

影響を受けるアプリケーションファイル名	AMI バージョンのブートストラップアクション	設定分類
emrfs-site.xml	configure-hadoop -e	emrfs-site
該当なし	s3get -s s3://custom-provider.jar -d /usr/share/aws/emr/auxlib/	emrfs-site (新しい設定 fs.s3.cse.encryptionMaterialsProvider.uri と併用)

全分類の一覧については、[アプリケーションの設定](#) を参照してください。

## アプリケーションの環境変数

AMI バージョンを使用する場合、hadoop-user-env.sh スクリプトが、configure-daemons ブートストラップアクションと組み合わせて使用され、Hadoop 環境を設定します。スクリプトには、以下のアクションが含まれます。

```
#!/bin/bash
export HADOOP_USER_CLASSPATH_FIRST=true;
echo "HADOOP_CLASSPATH=/path/to/my.jar" >> /home/hadoop/conf/hadoop-user-env.sh
```

Amazon EMR リリース 4.x では、次の例に示すように hadoop-env 設定分類を使用して同じ手順を実行します。

```
[
  {
    "Classification": "hadoop-env",
    "Properties": {
```



```

    },
    "Configurations":[
      {
        "Classification":"export",
        "Properties":{
          "HADOOP_USER_CLASSPATH_FIRST":"true",
          "HADOOP_CLASSPATH":"/path/to/my.jar"
        }
      }
    ]
  }
]

```

別の例として、`configure-daemons` の使用、`--namenode-heap-size=2048` を渡すこと、`--namenode-opts=-XX:GCTimeRatio=19` は以下の設定分類と同じです。

```

[
  {
    "Classification":"hadoop-env",
    "Properties":{

    },
    "Configurations":[
      {
        "Classification":"export",
        "Properties":{
          "HADOOP_DATANODE_HEAPSIZE": "2048",
          "HADOOP_NAMENODE_OPTS": "-XX:GCTimeRatio=19"
        }
      }
    ]
  }
]

```

他のアプリケーションの環境変数は、`/home/hadoop/.bashrc` では定義されなくなりました。代わりに、コンポーネントやアプリケーションごとに、主に `/etc/default` のような `/etc/default/hadoop` ファイルで設定されます。アプリケーション RPM によってインストールされる `/usr/bin/` にあるラッパースクリプトでも、バイナリスクリプトを実際に含める前に、追加の環境変数を設定できます。


## サービスポート

AMI バージョンを使用する場合は、一部のサービスでカスタムポートを使用します。

### ポート設定の変更箇所

設定	AMI バージョン 3.x	オープンソースのデフォルト
fs.default.name	hdfs://emrDeterminedIP:9000	デフォルト (hdfs:// <i>emrDeterminedIP</i> :8020)
dfs.datanode.address	0.0.0.0:9200	デフォルト (0.0.0.0:50010)
dfs.datanode.http.address	0.0.0.0:9102	デフォルト (0.0.0.0:50075)
dfs.datanode.https.address	0.0.0.0:9402	デフォルト (0.0.0.0:50475)
dfs.datanode.ipc.address	0.0.0.0:9201	デフォルト (0.0.0.0:50020)
dfs.http.address	0.0.0.0:9101	デフォルト (0.0.0.0:50070)
dfs.https.address	0.0.0.0:9202	デフォルト (0.0.0.0:50470)
dfs.secondary.http.address	0.0.0.0:9104	デフォルト (0.0.0.0:50090)
yarn.nodemanager.address	0.0.0.0:9103	デフォルト (\${yarn.nodemanager.hostname}:0)
yarn.nodemanager.localizer.address	0.0.0.0:9033	デフォルト (\${yarn.nodemanager.hostname}:8040)
yarn.nodemanager.webapp.address	0.0.0.0:9035	デフォルト (\${yarn.nodemanager.hostname}:8042)
yarn.resourcemanager.address	<i>emrDeterminedIP</i> :9022	デフォルト (\${yarn.resourcemanager.hostname}:8032)

設定	AMI バージョン 3.x	オープンソースのデフォルト
yarn.resourcemanager.admin.address	<i>emrDeterminedIP</i> :9025	デフォルト ( $\{\text{yarn.resourcemanager.hostname}\}$ :8033)
yarn.resourcemanager.resource-tracker.address	<i>emrDeterminedIP</i> :9023	デフォルト ( $\{\text{yarn.resourcemanager.hostname}\}$ :8031)
yarn.resourcemanager.scheduler.address	<i>emrDeterminedIP</i> :9024	デフォルト ( $\{\text{yarn.resourcemanager.hostname}\}$ :8030)
yarn.resourcemanager.webapp.address	0.0.0.0:9026	デフォルト ( $\{\text{yarn.resourcemanager.hostname}\}$ :8088)
yarn.web-proxy.address	<i>emrDeterminedIP</i> :9046	デフォルト (値なし)
yarn.resourcemanager.hostname	0.0.0.0 (デフォルト)	<i>emrDeterminedIP</i>

 Note

*emrDeterminedIP* は、Amazon EMR で生成される IP アドレスです。

## [ユーザー]

AMI バージョンを使用する場合、ユーザー `hadoop` がすべてのプロセスを実行し、すべてのファイルを保有します。Amazon EMR リリースバージョン 4.0.0 以降では、ユーザーはアプリケーションレベルやコンポーネントレベルで存在します。

## インストールシーケンス、インストールされているアーティファクト、およびログファイルの場所

AMI バージョンを使用する場合、アプリケーションのアーティファクトおよびその設定ディレクトリは、`/home/hadoop/application` ディレクトリにインストールされます。たとえば、Hive をインストールした場合は、ディレクトリは `/home/hadoop/hive` でした。Amazon EMR リリース 4.0.0 以降では、アプリケーションのアーティファクトは `/usr/lib/application` ディレクトリにインストールされます。AMI バージョンを使用する場合、ログファイルはさまざまな場所にあります。次の表に、場所を示します。

### Amazon S3 のログの場所の変更

デーモンまたはアプリケーション	ディレクトリの場所
instance-state	node/ <i>instance-id</i> /instance-state/
hadoop-hdfs-namenode	デーモン/ <i>##### ID</i> /hadoop-hadoop-namenode.log
hadoop-hdfs-datanode	デーモン/ <i>##### ID</i> /hadoop-hadoop-datanode.log
hadoop-yarn (ResourceManager )	デーモン/ <i>##### ID</i> /yarn-hadoop-resource-manager
hadoop-yarn (プロキシサーバー)	デーモン/ <i>##### ID</i> /yarn-hadoop-proxyserver
mapred-historyserver	daemons/ <i>instance-id</i> /
httpfs	daemons/ <i>instance-id</i> /httpfs.log
hive-server	node/ <i>instance-id</i> /hive-server/hive-server.log
hive-metastore	node/ <i>instance-id</i> /apps/hive.log
Hive CLI	node/ <i>instance-id</i> /apps/hive.log

デーモンまたはアプリケーション	ディレクトリの場所
YARN アプリケーションのユーザーログおよびコンテナログ	task-attempts/
Mahout	該当なし
Pig	該当なし
spark-historyserver	該当なし
MapReduce ジョブ履歴ファイル	jobs/

## Command runner

AMI バージョンを使用する場合、`/home/hadoop/contrib/streaming/hadoop-streaming.jar` のような多くのスクリプトまたはプログラムが、シェルのログインパス環境に置かれていないため、`command-runner.jar` または `script-runner.jar` などの jar ファイルを使用する際に完全パスを指定する必要があります。`command-runner.jar` が AMI に置かれるため、`script-runner.jar` を使用する場合のように完全 URI を指定する必要がなくなりました。

## レプリケーション係数

レプリケーション係数を使用すると、Hadoop JVM をいつ起動するかを設定できます。Hadoop JVM は、タスクごとに新しく起動できます。これにより、タスクの独立性を高めることも、タスク間で JVM を共有してフレームワークのオーバーヘッドを低下させることもできます。多くの小さいファイルを処理する場合は、JVM を何回も再利用して起動のコストを分散させることに意味があります。ただし、各タスクが長時間かかる場合や大量のデータを処理する場合は、後続のタスクのためにすべてのメモリが確実に解放されるように、JVM を再利用しない選択もできます。AMI のバージョンを使用する場合は、`configure-hadoop` ブートストラップアクションを使用してレプリケーション係数をカスタマイズすることができ、`mapred.job.reuse.jvm.num.tasks` プロパティを設定します。

次の例は、無限の JVM 再利用の JVM 再利用係数の設定を示しています。

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m3.xlarge \
InstanceGroupType=CORE,InstanceCount=2,InstanceType=m3.xlarge \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

Amazon EMR の以前の AMI バージョンの Hive アプリケーションの仕様。

## ログファイル

Amazon EMR AMI バージョン 2.x および 3.x を使用する場合、Hive ログは /mnt/var/log/apps/ に保存されます。Hive の複数のバージョンを同時にサポートするために、ログファイル名は次の表に示すように、実行する Hive のバージョンによって異なります。

Hive のバージョン	ログファイル名
0.13.1	hive.log
0.11.0	hive_0110.log

**Note**

このバージョンからは、Amazon EMR で、バージョンニング無効のファイル名 hive.log が使用されます。マイナーバージョンは、メジャーバージョンと同じログの場所を共有します。

Hive のバージョン	ログファイル名
	<p><b>Note</b></p> <p>0.11.0.1 のような Hive 0.11.0 のマイナーバージョンでは、ログファイルの場所は Hive 0.11.0 と同じです。</p>
0.8.1	<p>hive_081.log</p> <p><b>Note</b></p> <p>Hive 0.8.1.1 のような Hive 0.8.1 のマイナーバージョンでは、ログファイルの場所は Hive 0.8.1 と同じです。</p>
0.7.1	<p>hive_07_1.log</p> <p><b>Note</b></p> <p>Hive 0.7.1.3 や Hive 0.7.1.4 といった Hive 0.7.1 のマイナーバージョンでは、ログファイルの場所は Hive 0.7.1 と同じです。</p>
0.7	hive_07.log
0.5	hive_05.log
0.4	hive.log

## 分割入力機能

0.13.1 より前の Hive バージョン (Amazon EMR AMI バージョン 3.11.0 より前) を使用して分割入力機能を実装するには、次のようにします。

```
hive> set hive.input.format=org.apache.hadoop.hive.q1.io.HiveCombineSplitsInputFormat;
hive> set mapred.min.split.size=100000000;
```

この機能は Hive 0.13.1 で廃止されました。Amazon EMR AMI バージョン 3.11.0 で同じ分割入力形式の機能を得るには、以下を使用します。

```
set hive.hadoop.supports.plittable.combineinputformat=true;
```

## Thrift サービスポート

Thrift は RPC のフレームワークで、これによって定義されるコンパクトなバイナリシリアライズ形式を使用して、後から分析に使用するためのデータ構造を維持します。通常、Hive は次のポートで稼働するようにサーバーを設定します。

Hive のバージョン	ポート番号
Hive 0.13.1	10000
Hive 0.11.0	10004
Hive 0.8.1	10003
Hive 0.7.1	10002
Hive 0.7	10001
Hive 0.5	10000

thrift サービスの詳細については、<http://wiki.apache.org/thrift/> を参照してください。

## Hive を使用してパーティションを復元する

Amazon EMR の Hive クエリ言語には、Amazon S3 内にあるテーブルデータからテーブルのパーティションを復元するステートメントが含まれています。次の例でこれを示します。

```
CREATE EXTERNAL TABLE (json string) raw_impression
PARTITIONED BY (dt string)
LOCATION 's3://elastic-mapreduce/samples/hive-ads/tables/impressions';
ALTER TABLE logs RECOVER PARTITIONS;
```



パーティションのディレクトリおよびデータは、テーブル定義で指定された場所に格納し、名前は Hive の表記規則に従って、`dt=2009-01-01` のように記述する必要があります。

#### Note

Hive 0.13.1 以降では、この機能は `msck repair table` を使用してネイティブにサポートされるため、`recover partitions` はサポートされません。詳細については、<https://cwiki.apache.org/confluence/display/Hive/LanguageManual+DDL> を参照してください。

## Hive 変数をスクリプトに渡す

を使用して Hive ステップに変数を渡すには AWS CLI、次のコマンドを入力します。`myKey` を EC2 キーペアの名前に置き換え、`mybucket` をバケット名に置き換えます。この例で、`SAMPLE` は、先頭に `-d` スイッチが付いた変数の値です。この変数は Hive スクリプトで `${SAMPLE}` のように定義されます。

#### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.9 \  
--applications Name=Hue Name=Hive Name=Pig \  
--use-default-roles --ec2-attributes KeyName=myKey \  
--instance-type m3.xlarge --instance-count 3 \  
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,\  
Args=[-f,s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,\  
INPUT=s3://elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://mybucket/hive-ads/  
output/,\  
-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

## 外部メタストアの場所を指定する

次の手順では、Hive メタストアの場所として設定したデフォルト値を上書きし、新しく設定したメタストアの場所を使用してクラスターを起動する方法を示します。

## EMR クラスターの外にメタストアを作成するには

1. Amazon RDS を使用して MySQL または Aurora データベースを作成します。

Amazon RDS データベースの作成方法については、「[Amazon RDS の開始方法](#)」を参照してください。

2. セキュリティグループを変更して、データベースと ElasticMapReduce-Master セキュリティグループ間の JDBC 接続を許可します。

アクセス用のセキュリティグループの変更方法については、「Amazon RDS ユーザーガイド」の「[Amazon RDS セキュリティグループ](#)」を参照してください。

3. JDBC 設定値は、hive-site.xml に格納されています。
  - a. 以下を含む hive-site.xml 設定ファイルを作成します。

```
<configuration>
  <property>
    <name>javax.jdo.option.ConnectionURL</name>
    <value>jdbc:mariadb://hostname:3306/hive?createDatabaseIfNotExist=true</value>
    <description>JDBC connect string for a JDBC metastore</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionUserName</name>
    <value>hive</value>
    <description>Username to use against metastore database</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionPassword</name>
    <value>password</value>
    <description>Password to use against metastore database</description>
  </property>
</configuration>
```

*hostname* は、データベースを実行する Amazon RDS インスタンスの DNS アドレスです。*username* と *password* は、データベースの認証情報です。MySQL データベースインスタンスおよび Aurora データベースインスタンスへの接続の詳細については、「Amazon RDS ユーザーガイド」の「[MySQL データベースエンジンを実行している DB インスタンスへの接続](#)」および「[Athena DB クラスターへの接続](#)」を参照してください。

JDBC ドライバーは Amazon EMR によってインストールされます。

**Note**

値のプロパティには、スペースや改行を含めることはできません。すべてを 1 行にする必要があります。

- b. Amazon S3 上の `hive-site.xml` などの場所に `s3://mybucket/hive-site.xml` ファイルを保存します。
4. クラスターを作成し、カスタマイズされた `hive-site.xml` ファイルの Amazon S3 の場所を指定します。

次のコマンド例は、これを行う AWS CLI コマンドを示しています。

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.10 \  
--applications Name=Hue Name=Hive Name=Pig \  
--use-default-roles --ec2-attributes KeyName=myKey \  
--instance-type m3.xlarge --instance-count 3 \  
--bootstrap-actions Name="Install Hive Site Configuration",\  
Path="s3://region.elasticmapreduce/libs/hive/hive-script",\  
Args=["--base-path", "s3://elasticmapreduce/libs/hive", "--install-hive-site",\  
"--hive-site=s3://mybucket/hive-site.xml", "--hive-versions", "latest"]
```

## JDBC を使用して Hive に接続する

JDBC を介して Hive に接続するには、JDBC ドライバをダウンロードして、SQL クライアントをインストールする必要があります。以下の例は、SQL Workbench/J を使用して、JDBC を使用している Hive に接続する方法を示しています。

## JDBC ドライバをダウンロードするには

1. アクセスする Hive のバージョンに対応しているドライバをダウンロードし、展開します。Hive バージョンは、Amazon EMR クラスターを作成するときを選択した AMI に応じて異なります。
  - Hive 0.13.1 JDBC ドライバー: [https://amazon-odbc-jdbc-drivers.s3.amazonaws.com/public/AmazonHiveJDBC\\_1.0.4.1004.zip](https://amazon-odbc-jdbc-drivers.s3.amazonaws.com/public/AmazonHiveJDBC_1.0.4.1004.zip)
  - Hive 0.11.0 JDBC ドライバー: <https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.11.0>
  - Hive 0.8.1 JDBC ドライバー: <https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.8.1>
2. SQL Workbench/J をインストールします。詳細については、SQL Workbench/J ユーザーズマニュアルの「[Installing and starting SQL Workbench/J](#)」を参照してください。
3. クラスターマスターノードへの SSH トンネルを作成します。Hive のバージョンによって、接続するポートが異なります。Linux ssh ユーザーの場合は以下の表、Windows ユーザーの場合は PuTTY コマンドのコマンドの例が提供されています。

### Linux SSH コマンド

Hive のバージョン	Command
0.13.1	<code>ssh -o ServerAliveInterval=10 -i <i>path-to-key-file</i> -N -L 10000:localhost:10000 hadoop@ <i>master-public-dns-name</i></code>
0.11.0	<code>ssh -o ServerAliveInterval=10 -i <i>path-to-key-file</i> -N -L 10004:localhost:10004 hadoop@ <i>master-public-dns-name</i></code>
0.8.1	<code>ssh -o ServerAliveInterval=10 -i <i>path-to-key-file</i> -N -L 10003:localhost:10003 hadoop@ <i>master-public-dns-name</i></code>
0.7.1	<code>ssh -o ServerAliveInterval=10 -i <i>path-to-key-file</i> -N -L 10002:localhost:10002 hadoop@ <i>master-public-dns-name</i></code>
0.7	<code>ssh -o ServerAliveInterval=10 -i <i>path-to-key-file</i> -N -L 10001:localhost:10001 hadoop@ <i>master-public-dns-name</i></code>

Hive のバージョン	Command
0.5	<code>ssh -o ServerAliveInterval=10 -i <i>path-to-key-file</i> -N -L 10000:localhost:10000 hadoop@ <i>master-public-dns-name</i></code>

#### Windows PuTTY のトンネルの設定

Hive のバージョン	トンネルの設定
0.13.1	送信元ポート : 10000 送信先 : <i>master-public-dns-name</i> : 10000
0.11.0	送信元ポート : 10004 送信先 : <i>master-public-dns-name</i> :10004
0.8.1	送信元ポート : 10003 送信先 : <i>master-public-dns-name</i> :10003

#### 4. SQL Workbench に JDBC ドライバを追加します。

- [Select Connection Profile] (接続プロファイルの選択) ダイアログボックスで [Manage Drivers] (ドライバーの管理) を選択します。
- [Create a new entry] (新しいエントリーの作成) (空白のページ) アイコンを選択します。
- 名前フィールドに、**Hive JDBC** と入力します。
- [Library] で [Select the JAR file(s)] アイコンをクリックします。
- 次の表に示すように JAR ファイルを選択します。

Hive ドライババージョン	追加する JAR ファイル
0.13.1	<pre>hive_metastore.jar hive_service.jar HiveJDBC3.jar libfb303-0.9.0.jar libthrift-0.9.0.jar</pre>

Hive ドライババージョン	追加する JAR ファイル
	<pre>log4j-1.2.14.jar ql.jar slf4j-api-1.5.8.jar slf4j-log4j12-1.5.8.jar TCLIServiceClient.jar</pre>
0.11.0	<pre>hadoop-core-1.0.3.jar hive-exec-0.11.0.jar hive-jdbc-0.11.0.jar hive-metastore-0.11.0.jar hive-service-0.11.0.jar libfb303-0.9.0.jar commons-logging-1.0.4.jar slf4j-api-1.6.1.jar</pre>
0.8.1	<pre>hadoop-core-0.20.205.jar hive-exec-0.8.1.jar hive-jdbc-0.8.1.jar hive-metastore-0.8.1.jar hive-service-0.8.1.jar libfb303-0.7.0.jar libthrift-0.7.0.jar log4j-1.2.15.jar slf4j-api-1.6.1.jar slf4j-log4j12-1.6.1.jar</pre>
0.7.1	<pre>hadoop-0.20-core.jar hive-exec-0.7.1.jar hive-jdbc-0.7.1.jar hive-metastore-0.7.1.jar hive-service-0.7.1.jar libfb303.jar commons-logging-1.0.4.jar slf4j-api-1.6.1.jar slf4j-log4j12-1.6.1.jar</pre>

Hive ドライババージョン	追加する JAR ファイル
0.7	<pre> hadoop-0.20-core.jar hive-exec-0.7.0.jar hive-jdbc-0.7.0.jar hive-metastore-0.7.0.jar hive-service-0.7.0.jar libfb303.jar commons-logging-1.0.4.jar slf4j-api-1.5.6.jar slf4j-log4j12-1.5.6.jar </pre>
0.5	<pre> hadoop-0.20-core.jar hive-exec-0.5.0.jar hive-jdbc-0.5.0.jar hive-metastore-0.5.0.jar hive-service-0.5.0.jar libfb303.jar log4j-1.2.15.jar commons-logging-1.0.4.jar </pre>

- f. [Please select one driver] (ドライバーを 1 つ選択してください) ダイアログボックスで次の表に従ってドライバを選択し、[OK] をクリックします。

Hive のバージョン	ドライバのクラス名
0.13.1	<pre>com.amazon.hive.jdbc3.HS2Driver</pre>
0.11.0	<pre>org.apache.hadoop.hive.jdbc.HiveDriver.jar</pre>
0.8.1	<pre>org.apache.hadoop.hive.jdbc.HiveDriver.jar</pre>

Hive のバージョン	ドライバのクラス名
0.7.1	<code>org.apache.hadoop.hive.jdbc.HiveDriver.jar</code>
0.7	<code>org.apache.hadoop.hive.jdbc.HiveDriver.jar</code>
0.5	<code>org.apache.hadoop.hive.jdbc.HiveDriver.jar</code>

5. [Select Connection Profile] (接続プロファイルの選択) ダイアログボックスに戻ったら、[ドライバー] フィールドが [Hive JDBC] に設定されていることを確認し、次の表に従って [URL] フィールドに JDBC 接続文字列を入力します。

Hive のバージョン	JDBC 接続文字列
0.13.1	<code>jdbc:hive2://localhost:10000/default</code>
0.11.0	<code>jdbc:hive://localhost:10004/default</code>
0.8.1	<code>jdbc:hive://localhost:10003/default</code>

クラスターが AMI バージョン 3.3.1 以降を使用している場合は、[Select Connection Profile] (接続プロファイルの選択) ダイアログボックスの [ユーザー名] フィールドに **hadoop** と入力します。



## Amazon EMR の以前の AMI バージョンの HBase アプリケーションの仕様

### サポートされる HBase バージョン

HBase バージョン	AMI のバージョン	AWS CLI 設定パラメータ	HBase バージョンの詳細
<a href="#">0.94.18</a>	3.1.0 以降	<pre>--ami-version 3.1  --ami-version 3.2  --ami-version 3.3  --applications Name=HBase</pre>	<ul style="list-style-type: none"> <li>バグの修正と機能強化。</li> </ul>
<a href="#">0.94.7</a>	3.0-3.0.4	<pre>--ami-version 3.0  --applications Name=HBase</pre>	
<a href="#">0.92</a>	2.2 以降	<pre>--ami-version 2.2 or later  --applications Name=HBase</pre>	

### HBase クラスターの前提条件

Amazon EMR AMI バージョン 2.x および 3.x を使用して作成されたクラスターは、HBase の次の要件を満たす必要があります。

- AWS CLI (オプション) - コマンドラインを使用して HBase を操作するには、の最新バージョンをダウンロードしてインストールします AWS CLI。詳細については、「AWS Command Line

Interface ユーザーガイド」の「[AWS Command Line Interfaceのインストール](#)」を参照してください。

- 2 つ以上のインスタンス (オプション) - クラスターのマスターノードでは HBase マスターサーバーと Zookeeper、タスクノードでは HBase リージョンサーバーが実行されます。最大のパフォーマンスを確保するには、HBase クラスターが 2 つ以上の EC2 インスタンスを実行する必要がありますが、評価目的の場合は HBase を単一ノードで実行できます。
- 長時間稼働クラスター - HBase は長時間稼働クラスター上でのみ実行されます。デフォルトでは、CLI および Amazon EMR コンソールは長時間稼働クラスターを作成します。
- Amazon EC2 キーペア (推奨) - マスターノードと接続して HBase シェルコマンドを実行するために Secure Shell (SSH) ネットワークプロトコルを使用するには、クラスターを作成するときに Amazon EC2 キーペアを使用する必要があります。
- 正しい AMI バージョンおよび Hadoop バージョン - HBase クラスターは、現在 Hadoop 20.205 以降でのみサポートされています。
- Ganglia (オプション) - HBase パフォーマンスメトリクスをモニタリングするには、クラスターの作成時に Ganglia をインストールします。
- ログ用 Amazon S3 バケット (オプション) - HBase 用ログはマスターノードで使用できます。これらのログを Amazon S3 にコピーするには、クラスターの作成時にログファイルを受け取るための S3 バケットを指定します。

## HBase を使用するクラスターの作成

次の表に、コンソールを使用し、Amazon EMR AMI リリースバージョンを使用して HBase を使用するクラスターを作成する際に利用できるオプションを示します。

フィールド	アクション
Restore from backup	HBase クラスターに、Amazon S3 に保存されているデータを事前にロードするかどうかを指定します。
Backup location	Amazon S3 で復元するバックアップがある場所の URI を指定します。
Backup version	[Backup Location] で指定した場所にある、使用するバックアップのバージョン名を指定します ( オプション )。このフィールドに何も指定しない場合は、[バックアップの場所] で指定した場所にある最新のバックアップが、Amazon EMR によって新しい HBase クラスターに設定されます。

フィールド	アクション
Schedule Regular Backups	自動増分バックアップをスケジュールするかどうかを指定します。最初は完全バックアップによって、以降の差分バックアップのベースラインとなるバックアップが作成されます。
Consistent backup	バックアップの整合性を確保するかどうかを指定します。バックアップの整合性とは、バックアップ処理の初期段階であるノード同期中に、書き込みオペレーションを一時停止することです。つまり、一時停止された書き込みオペレーションはすべてキューに置かれ、同期が完了したら再開されます。
Backup frequency	スケジュールされたバックアップの頻度 (日/時間/分単位)。
Backup location	バックアップが保存される Amazon S3 の URI。差分バックアップの正確性を確保するために、バックアップは、HBase クラスターごとにそれぞれ異なる場所に格納する必要があります。
Backup start time	最初のバックアップの時間を指定します。これを now に設定した場合、クラスターが実行されると直ちに最初のバックアップが行われます。または、日付と時刻を <a href="#">ISO 形式</a> で入力します。たとえば、2012-06-15T20:00Z を入力すると、2012 年 6 月 15 日午後 8 時 (UTC) に設定されます。

次の AWS CLI コマンド例では、HBase およびその他のアプリケーションを使用してクラスターを起動します。

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
  --applications Name=Hue Name=Hive Name=Pig Name=HBase \
  --use-default-roles --ec2-attributes KeyName=myKey \
  --instance-type c1.xlarge --instance-count 3 --termination-protected
```

Hive クラスターと HBase クラスターの間の接続が確立されたら ( 前述の手順を参照 )、Hive に外部テーブルを作成することで、HBase クラスターに格納されているデータにアクセスできます。

次の例では、Hive プロンプトからの実行時に、inputTable という HBase テーブルに格納されているデータを参照する外部テーブルを作成しています。その後、Hive ステートメントで inputTable を参照し、HBase クラスターに格納されているデータに対してクエリを実行したり、そのデータを変更したりできます。

#### Note

以下の例では、AMI 2.3.3 で protobuf-java-2.4.0a.jar を使用しますが、サンプルをご自分のバージョンに合わせてお使いください。お持ちの Protocol Buffers の JAR のバージョンをチェックするには、Hive コマンドプロンプトでコマンドを実行します: `! ls /home/hadoop/lib;`

```
add jar lib/emr-metrics-1.0.jar ;
    add jar lib/protobuf-java-2.4.0a.jar ;

set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com ;

create external table inputTable (key string, value string)
    stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
    with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
    tblproperties ("hbase.table.name" = "t1");

select count(*) from inputTable ;
```

## HBase 設定のカスタマイズ

デフォルトの設定はほとんどのアプリケーションで動作しますが、HBase 設定を変更する柔軟性も確保されています。これを行うには、次の 2 つのブートストラップアクションスクリプトのいずれかを実行します。

- `configure-hbase-daemons`— マスター、リージョンサーバー、およびゾーンキーパーデーモンのプロパティを設定します。これらのプロパティには、HBase デーモンの開始時に Java Virtual Machine ( JVM ) に渡すヒープサイズおよびオプションが含まれます。これらのプロパティは、ブートストラップアクションの引数として設定します。このブートストラップアクションは、HBase クラスターの `/home/hadoop/conf/hbase-user-env.sh` 設定ファイルを変更します。

- `configure-hbase` - HBase マスターのバインド先ポート、CLI クライアントがアクションを再試行する最大回数など、HBase サイト固有の設定を指定します。これらの をブートストラップアクションの引数 `one-by-one` として設定することも、Amazon S3 の XML 設定ファイルの場所を指定することもできます。このブートストラップアクションでは、HBase クラスターの `/home/hadoop/conf/hbase-site.xml` 設定ファイルが変更されます。

### Note

これらのスクリプトを実行できるのは、他のブートストラップアクションの場合と同様、クラスターの作成時のみです。このスクリプトを使用して、現在実行中の HBase クラスターの設定を変更することはできません。

`configure-hbase` または `configure-hbase-daemons` ブートストラップアクションを実行すると、指定した値がデフォルト値を上書きします。明示的に設定しなかった値には、デフォルト値が設定されます。

これらのブートストラップアクションでの HBase の設定操作は、Amazon EMR でブートストラップアクションを使用して Hadoop 設定と Hadoop デーモンのプロパティを設定するのと似ています。異なるのは、HBase にはプロセスごとのメモリオプションがないという点です。メモリオプションは、`--daemon-opts` 引数を使用して設定されます。`daemon` は、設定するデーモンの名前に置き換えてください。

## HBase デーモンを設定する

Amazon EMR には、HBase デーモンの設定変更に使用できるブートストラップアクション `s3://region.elasticmapreduce/bootstrap-actions/configure-hbase-daemons` が用意されています。`region` は、HBase クラスターを起動するリージョンです。

を使用して HBase デーモンを設定するには AWS CLI、クラスターを起動するときに `configure-hbase-daemons` ブートストラップアクションを追加して、1 つ以上の HBase デーモンを設定します。以下のプロパティを設定できます。

プロパティ	説明
<code>hbase-master-opts</code>	JVM によるマスターデーモンの実行方法をコントロールするオプション。設定されている場合は、これらによ

プロパティ	説明
	てデフォルトの HBASE_MASTER_OPTS 変数が上書きされます。
regionserver-opts	JVM によるリージョンサーバーデーモンの実行方法をコントロールするオプション。設定されている場合は、これらによってデフォルトの HBASE_REGIONSERVER_OPTS 変数が上書きされます。
zookeeper-opts	JVM による zookeeper デーモンの実行方法をコントロールするオプション。設定されている場合は、これらによってデフォルトの HBASE_ZOOKEEPER_OPTS 変数が上書きされます。

これらのオプションの詳細については、HBase ドキュメントの「[hbase-env.sh](#)」を参照してください。

zookeeper-opts および hbase-master-opts の値を設定するためのブートストラップアクションを次の例に示します。

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-
daemons,\
Args=["--hbase-zookeeper-opts=-Xmx1024m -XX:GCTimeRatio=19", "--hbase-master-opts=-
Xmx2048m", "--hbase-regionserver-opts=-Xmx4096m"]
```

## HBase サイト設定を指定する

Amazon EMR には、HBase の設定変更を使用できるブートストラップアクション `s3://elasticmapreduce/bootstrap-actions/configure-hbase` が用意されています。設定値をブートストラップアクションの引数 `one-by-one` として設定することも、Amazon S3 の XML 設定ファイルの場所を指定することもできます。設定値の設定 `one-by-one` は、少数の設定のみが必要な場合に便利です。追加する変更が多数ある場合、または再利用できるように設定を保存したい場合は、XML ファイルを使用して設定することをお勧めします。

### Note

Amazon S3 バケット名には、`s3://region.elasticmapreduce/bootstrap-actions/configure-hbase` のようにリージョンプレフィックスを付けることができます。*region* は、HBase クラスターを起動するリージョンです。

このブートストラップアクションでは、HBase クラスターの `/home/hadoop/conf/hbase-site.xml` 設定ファイルが変更されます。ブートストラップアクションは、HBase クラスターの起動時にのみ実行できます。

設定可能な HBase サイト設定の詳細については、HBase のドキュメントの「[Default configuration](#)」を参照してください。

HBase クラスターの起動時に `configure-hbase` ブートストラップアクションを設定し、`hbase-site.xml` で変更する値を指定します。

を使用して個々の HBase サイト設定を指定するには AWS CLI

- `hbase.hregion.max.filesize` 設定を変更するには、次のコマンドを入力し、*myKey* を Amazon EC2 キーペアの名前に置き換えます。

### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \  
  \
```

```
--applications Name=Hue Name=Hive Name=Pig Name=HBase \  
--use-default-roles --ec2-attributes KeyName=myKey \  
--instance-type c1.xlarge --instance-count 3 --termination-protected \  
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-  
hbase,Args=["-s", "hbase.hregion.max.filesize=52428800"]
```

を使用して XML ファイルで HBase サイト設定を指定するには AWS CLI

1. カスタムバージョンの `hbase-site.xml` を作成します。カスタムファイルは有効な XML である必要があります。エラーが発生する可能性を少なくするには、`/home/hadoop/conf/hbase-site.xml` の Amazon EMR HBase マスターノードにある `hbase-site.xml` のデフォルトのコピーから開始し、そのファイルのコピーを編集します。ファイルをゼロから作成する必要はありません。新しいファイルに新しい名前を付けるか、`hbase-site.xml` をそのまま使用します。
2. カスタム `hbase-site.xml` ファイルを Amazon S3 バケットにアップロードします。クラスターを起動する AWS アカウントが ファイルにアクセスできるように、アクセス許可が設定されている必要があります。クラスターを起動する AWS アカウントが Amazon S3 バケットも所有している場合、そのアカウントは にアクセスできます。
3. HBase クラスターの起動時に `configure-hbase` ブートストラップアクションを設定し、カスタム `hbase-site.xml` ファイルの場所を含めます。次の例では、HBase サイト設定の値を、`s3://mybucket/my-hbase-site.xml` で指定した値に設定しています。次のコマンドを入力し、*myKey* を EC2 キーペアの名前に置き換え、*mybucket* を Amazon S3 バケットの名前に置き換えます。

#### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \  
  --applications Name=Hue Name=Hive Name=Pig Name=HBase \  
  --use-default-roles --ec2-attributes KeyName=myKey \  
  --instance-type c1.xlarge --instance-count 3 --termination-protected \  
  --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-  
hbase,Args=["--site-config-file", "s3://mybucket/config.xml"]
```



複数のオプションを指定して HBase オペレーションをカスタマイズする場合は、次の例に示すように、各キーと値のペアの前に `-s` オプションスイッチを追加する必要があります。

```
--bootstrap-actions s3://elasticmapreduce/bootstrap-actions/configure-hbase,Args=["-s","zookeeper.session.timeout=60000"]
```

プロキシを設定し、SSH 接続を開いた状態で、`http://master-public-dns-name:60010/master-status` でブラウザウィンドウを開くことで HBase UI を表示できます。ここで、*master-public-dns-name* は HBase クラスター内のマスターノードのパブリック DNS アドレスです。

現在の HBase ログを表示するには、SSH を使用してマスターノードに接続し、`mnt/var/log/hbase` ディレクトリに移動します。これらのログは、クラスターの起動時に Amazon S3 へのログを有効にしない限り、クラスター終了後に使用できなくなります。

## HBase をバックアップおよび復元する

Amazon EMR には、手動または自動化されたスケジュールに従って、HBase データを Amazon S3 にバックアップする機能が用意されています。完全バックアップと増分バックアップの両方を実行できます。HBase データのバックアップバージョンを作成した後、そのバージョンを HBase クラスターに復元できます。復元は現在実行中の HBase クラスターに対して行うことができます。また、バックアップされたデータが設定されたクラスターを新しく作成することもできます。

バックアップ処理中も、HBase は書き込みコマンドを実行し続けます。これにより、バックアップ中もクラスターを使用し続けることができますが、バックアップ中のデータと、並行して実行されている書き込みオペレーションとの間に不整合が発生するリスクがあります。発生する可能性がある不整合について理解するには、HBase では、書き込みオペレーションがクラスター内のノード全体に分散されるという点を考慮する必要があります。特定のノードがポーリングされた後、書き込みオペレーションが発生すると、そのデータはバックアップアーカイブに保存されません。HBase クラスターへの以前の書き込み（既にポーリングされたノードに送信された書き込み）がバックアップアーカイブに含まれない可能性がある一方で、それより後の書き込み（ポーリングされる前のノードに送信された書き込み）がアーカイブに保存されることもあります。

整合性のあるバックアップが必要な場合は、バックアップ処理の初期段階であるノード同期中に、HBase への書き込みを一時停止する必要があります。これを行うには、バックアップをリクエストするときに `--consistent` パラメータを指定します。このパラメータが指定されている場合は、この期間中の書き込みがキューに入れられ、同期が完了するとすぐに実行されます。定期バック

アップをスケジュールすることもできます。これにより徐々に発生する不整合が解決されます。あるバックアップパスで逃したデータが、次のパスでバックアップされるからです。

HBase データをバックアップするときに、クラスターごとにさまざまなバックアップディレクトリを指定する必要があります。これを簡単に行うには、クラスター識別子を、バックアップディレクトリに指定されたパスの一部として使用します。例えば `s3://mybucket/backups/j-3AEXXXXXX16F2` です。これにより、以降のすべての差分バックアップが適切な HBase クラスターを参照するようになります。

不要になった古いバックアップファイルを削除する準備ができたなら、まず HBase データの完全バックアップを行うことをお勧めします。これによりすべてのデータが保存されます。このデータは、その後の差分バックアップのベースとして使用できます。完全バックアップが完了した後、バックアップの場所に移動し、古いバックアップファイルを手動で削除できます。

HBase バックアッププロセスでは、コピーオペレーションに S3DistCp を使用します。これには、一時ファイルストレージ領域に関する特定の制限があります。

#### コンソールを使用して HBase をバックアップおよび復元する

コンソールを使用すると、新しいクラスターを起動して、以前の HBase バックアップのデータを入力することができます。また、HBase データの定期差分バックアップをスケジュールすることもできます。追加のバックアップおよび復元機能、たとえば、既に実行中のクラスターにデータを復元したり、手動バックアップを行ったり、自動完全バックアップをスケジュールしたりする機能は、CLI で使用できます。

コンソールを使用して、アーカイブされた HBase データを新しいクラスターに指定するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [クラスターを作成] を選択します。
3. [Software Configuration] セクションの [Additional Applications] で、[HBase] と [Configure and add] を選択します。
4. [Add Application] ダイアログボックスで、[Restore From Backup] チェックボックスをオンにします。
5. [Backup Location] (バックアップの場所) で、新しい HBase クラスターにロードするバックアップの場所を指定します。ここには、Amazon S3 URL を `s3://myawsbucket/backups/` という形式で指定する必要があります。

6. [Backup Version] では、値を設定して、ロードするバックアップバージョンの名前を指定するオプションがあります。[バックアップバージョン] に値を設定しない場合は、指定した場所にある最新のバックアップが Amazon EMR によってロードされます。
7. [Add] を選択して、必要に応じたオプションによるクラスターの作成に進みます。

コンソールを使用して HBase データの自動バックアップをスケジュールするには

1. [Software Configuration] セクションの [Additional Applications] で、[HBase] と [Configure and add] を選択します。
2. [Schedule Regular Backups] を選択します。
3. バックアップの整合性を確保するかどうかを指定します。バックアップの整合性とは、バックアップ処理の初期段階であるノード同期中に、書き込みオペレーションを一時停止することです。つまり、一時停止された書き込みオペレーションはすべてキューに置かれ、同期が完了したら再開されます。
4. バックアップの頻度を設定します。それには、[Backup Frequency] に数値を入力し、[Days]、[Hours]、または [Minutes] を選択します。最初に行われる自動バックアップは完全バックアップです。その後、指定したスケジュールに基づき、差分バックアップが Amazon EMR によって保存されます。
5. バックアップを保存する Amazon S3 内の場所を指定します。差分バックアップが適切に判断されるように、HBase クラスターをそれぞれ Amazon S3 内の別々の場所にバックアップする必要があります。
6. [Backup Start Time] で、バックアップを最初に行うタイミングの値を指定します。これを now に設定した場合、クラスターが実行されると直ちに最初のバックアップが行われます。または、日付と時刻を [ISO 形式](#) で入力します。たとえば、「2013-09-26T20:00Z」と入力すると、開始時間が 2013 年 9 月 26 日午後 8 時 (UTC) に設定されます。
7. [追加] を選択します。
8. 必要に応じて他のオプションを指定し、クラスターの作成を進めます。

## で HBase をモニタリングする CloudWatch

Amazon EMR は CloudWatch、HBase バックアップのモニタリングに使用できる 3 つのメトリクスをレポートします。これらのメトリクスは 5 分間隔で CloudWatch にプッシュされ、無料で提供されます。

メトリクス	説明
HBaseBackupFailed	<p>前回のバックアップに失敗したかどうか。デフォルトでは 0 に設定されており、前回のバックアップが失敗した場合は 1 に更新されます。このメトリクスは、HBase クラスターについてのみ報告されます。</p> <p>ユースケース: HBase バックアップを監視する</p> <p>単位: Count</p>
HBaseMostRecentBackupDuration	<p>前回のバックアップが完了するのにかかった時間。このメトリクスは、前回完了したバックアップが成功したか失敗したかに関係なく設定され、このメトリクスは、バックアップの進行中に、バックアップが開始してからの分数を返します。このメトリクスは、HBase クラスターについてのみ報告されます。</p> <p>ユースケース: HBase バックアップを監視する</p> <p>単位: Minutes</p>
HBaseTimeSinceLastSuccessfulBackup	<p>クラスターで前回成功した HBase バックアップが開始してから経過した分数。このメトリクスは、HBase クラスターについてのみ報告されます。</p> <p>ユースケース: HBase バックアップを監視する</p> <p>単位: Minutes</p>

## HBase の Ganglia の設定

configure-hbase-for-ganglia ブートストラップアクションを使用して HBase の Ganglia を設定します。このブートストラップアクションは、Ganglia にメトリクスを発行するように HBase を設定します。

クラスターを起動するときは、HBase と Ganglia を設定する必要があります。Ganglia のレポートを、実行中のクラスターに追加することはできません。

Ganglia では、/mnt/var/log/ganglia/rrds のサーバーにもログファイルを格納します。ログファイルが Amazon S3 バケットに保存されるようにクラスターを設定した場合は、Ganglia ログファイルも同様に同じ場所に保存されます。

Ganglia for HBase でクラスターを起動するには、次の例に示すようにconfigure-hbase-for-gangliaブートストラップアクションを使用します。

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \  
--applications Name=Hue Name=Hive Name=Pig Name=HBase Name=Ganglia \  
--use-default-roles --ec2-attributes KeyName=myKey \  
--instance-type c1.xlarge --instance-count 3 --termination-protected \  
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-for-ganglia
```

Ganglia を設定してクラスターを起動した後は、マスターノードで実行中のグラフィカルインターフェイスを使用して、Ganglia のグラフやレポートにアクセスできます。

## Amazon EMR の以前の AMI バージョンの Pig アプリケーションの仕様。

### サポートされている Pig のバージョン

クラスターに追加できる Pig のバージョンは、使用している Amazon EMR AMI のバージョンと Hadoop のバージョンによって決まります。次の表に、各バージョンの Pig と互換性のある AMI のバージョンおよび Hadoop のバージョンを示します。パフォーマンスの向上と新機能のメリットを受けられるように、最新バージョンの Pig を使用することをお勧めします。

API を使用して Pig をインストールする場合、 の呼び出し中に Pig をクラスターにロードするステップの引数--pig-versionsとして を指定しない限り、デフォルトバージョンが使用されます [RunJobFlow](#)。

Pig のバージョン	AMI のバージョン	設定パラメータ	Pig のバージョンの詳細
0.12.0 <a href="#">リリースノート</a> <a href="#">ドキュメント</a>	3.1.0 以降	--ami-version 3.1  --ami-version 3.2  --ami-version 3.3	以下のサポートが追加されています。  <ul style="list-style-type: none"> <li>JVM 実装を使用しない UDF のストリーミング</li> <li>ASSERT 演算子と IN 演算子</li> <li>CASE 式</li> <li>AvroStorage Pig 組み込み関数としての。</li> <li>ParquetLoader 組み込み関数 ParquetStorer としてのおよび</li> <li>BigInteger および BigDecimal タイプ</li> </ul>
0.11.1.1 <a href="#">リリースノート</a> <a href="#">ドキュメント</a>	2.2 以降	--pig-versions 0.11.1.1  --ami-version 2.2	入力が Amazon S3 に存在する PigStorage 場合の LOAD コマンドのパフォーマンスが向上します。
0.11.1 <a href="#">リリースノート</a> <a href="#">ドキュメント</a>	2.2 以降	--pig-versions 0.11.1  --ami-version 2.2	JDK 7、Hadoop 2、Groovy ユーザー定義関数、SchemaTuple 最適化、新しい演算子などのサポートが追加されました。詳細については、「 <a href="#">Pig</a>

Pig のバージョン	AMI のバージョン	設定パラメータ	Pig のバージョンの詳細
			<a href="#">0.11.1 change log</a> 」を参照してください。
0.9.2.2 <a href="#">リリースノート</a> <a href="#">ドキュメント</a>	2.2 以降	--pig-versions 0.9.2.2  --ami-version 2.2	Hadoop 1.0.3 のサポートを追加。
0.9.2.1 <a href="#">リリースノート</a> <a href="#">ドキュメント</a>	2.2 以降	--pig-versions 0.9.2.1  --ami-version 2.2	MapR のサポートを追加。
0.9.2 <a href="#">リリースノート</a> <a href="#">ドキュメント</a>	2.2 以降	--pig-versions 0.9.2  --ami-version 2.2	いくつかのパフォーマンス向上とバグ修正が行われています。Pig 0.9.2 での変更に関する情報については、「 <a href="#">Pig 0.9.2 change log</a> 」を参照してください。
0.9.1 <a href="#">リリースノート</a> <a href="#">ドキュメント</a>	2.0	--pig-versions 0.9.1  --ami-version 2.0	
0.6 <a href="#">リリースノート</a>	1.0	--pig-versions 0.6  --ami-version 1.0	

Pig のバージョン	AMI のバージョン	設定パラメータ	Pig のバージョンの詳細
0.3 <a href="#">リリースノート</a>	1.0	--pig-versions 0.3  --ami-version 1.0	

## Pig のバージョンの詳細

Amazon EMR では、追加の Amazon EMR パッチが適用される場合がある特定の Pig リリースをサポートします。Amazon EMR クラスターで実行する Pig のバージョンを設定することができます。これを行う方法については、「[Apache Pig](#)」を参照してください。以下のセクションでは、Pig の各バージョン、および Amazon EMR にロードされたバージョンに適用されるパッチについて説明します。

### Pig 向けパッチ

このセクションでは、Amazon EMR で使用可能な Pig バージョンに適用されるカスタムパッチについて説明します。

#### Pig 0.11.1.1 向けパッチ

Pig 0.11.1.1 の Amazon EMR バージョンは、入力が Amazon S3 に存在する PigStorage 場合、で LOAD コマンドのパフォーマンスを向上させるメンテナンスリリースです。

#### Pig 0.11.1 向けパッチ

Amazon EMR バージョンの Pig 0.11.1 には、Apache Software Foundation が提供するすべての更新と Pig バージョン 0.9.2.2 からの累積 Amazon EMR パッチが含まれます。ただし、Pig 0.11.1 には Amazon EMR 固有の新しいパッチはありません。

#### Pig 0.9.2 向けパッチ

Apache Pig 0.9.2 は、Pig のメンテナンスリリースです。Amazon EMR チームは今までに、Amazon EMR バージョンの Pig 0.9.2 に以下のパッチを適用しています。



パッチ	説明
PIG-1429	<p>Pig に Boolean データタイプを第 1 クラスデータタイプとして追加。詳細については、<a href="https://issues.apache.org/jira/browse/PIG-1429">https://issues.apache.org/jira/browse/PIG-1429</a> を参照してください。</p> <p>ステータス: コミット済み</p> <p>Fixed in Apache Pig Version: 0.10</p>
PIG-1824	<p>Jython UDF でインポートモジュールをサポート。詳細については、<a href="https://issues.apache.org/jira/browse/PIG-1824">https://issues.apache.org/jira/browse/PIG-1824</a> を参照してください。</p> <p>ステータス: コミット済み</p> <p>Fixed in Apache Pig Version: 0.10</p>
PIG-2010	<p>登録済みの JAR を分散キャッシュにバンドル。詳細については、<a href="https://issues.apache.org/jira/browse/PIG-2010">https://issues.apache.org/jira/browse/PIG-2010</a> を参照してください。</p> <p>ステータス: コミット済み</p> <p>Fixed in Apache Pig Version: 0.11</p>
PIG-2456	<p>ユーザーがデフォルトの Pig ステートメントを指定できる ~/.pigbootup ファイルを追加。詳細については、<a href="https://issues.apache.org/jira/browse/PIG-2456">https://issues.apache.org/jira/browse/PIG-2456</a> を参照してください。</p> <p>ステータス: コミット済み</p> <p>Fixed in Apache Pig Version: 0.11</p>
PIG-2623	<p>Amazon S3 パスを使用した UDF の登録をサポート。詳細については、<a href="https://issues.apache.org/jira/browse/PIG-2623">https://issues.apache.org/jira/browse/PIG-2623</a> を参照してください。</p> <p>ステータス: コミット済み</p>

パッチ	説明
	修正対象の Apache Pig のバージョン: 0.10、0.11

### Pig 0.9.1 向けパッチ

Amazon EMR チームは今までに、Amazon EMR バージョンの Pig 0.9.1 に以下のパッチを適用しています。

パッチ	説明
dfs での JAR ファイルと Pig スクリプトのサポート	<p>HDFS、Amazon S3 などの分散ファイルシステムに格納されているスクリプトの実行および JAR ファイルの登録をサポートする機能追加。詳細については、<a href="https://issues.apache.org/jira/browse/PIG-1505">https://issues.apache.org/jira/browse/PIG-1505</a> を参照してください。</p> <p>ステータス: コミット済み</p> <p>Fixed in Apache Pig Version: 0.8.0</p>
Pig における複数のファイルシステムのサポート	<p>あるファイルシステムから読み取ったデータを別のファイルシステムに書き込む Pig スクリプトに対するサポートを追加。詳細については、<a href="https://issues.apache.org/jira/browse/PIG-1564">https://issues.apache.org/jira/browse/PIG-1564</a> を参照してください。</p> <p>ステータス: 未コミット</p> <p>修正対象の Apache Pig バージョン: n/a</p>
Piggybank の datetime および string UDF の追加	<p>カスタムの Pig スクリプトをサポートするために datetime および string UDF を追加。詳細については、<a href="https://issues.apache.org/jira/browse/PIG-1565">https://issues.apache.org/jira/browse/PIG-1565</a> を参照してください。</p> <p>ステータス: 未コミット</p> <p>修正対象の Apache Pig バージョン: n/a</p>

## インタラクティブおよびバッチ Pig クラスタ

Amazon EMR では、Pig のスクリプトを次の 2 つのモードで実行できます。

- インタラクティブ
- バッチ

コンソールまたは を使用して長時間実行されるクラスタを起動する場合 AWS CLI、ssh を Hadoop ユーザーとしてマスターノードに接続し、Grunt シェルを使用して Pig スクリプトをインタラクティブに開発および実行できます。Pig をインタラクティブに使用すると、バッチモードよりも簡単に Pig スクリプトを変更できます。インタラクティブモードで Pig スクリプトの変更が完了したら、そのスクリプトを Amazon S3 にアップロードし、バッチモードを使用してスクリプトを本稼働環境で実行できます。また、実行中のクラスタで Pig コマンドをインタラクティブに送信し、必要に応じてデータの分析や変換を行うことができます。

バッチモードでは、Pig スクリプトを Amazon S3 にアップロードし、クラスタにステップとして作業を送信します。Pig ステップは、長時間稼働クラスタまたは一時的なクラスタに送信できません。

## Amazon EMR の以前の AMI バージョンの Spark アプリケーションの様

### Spark のインタラクティブな使用またはバッチモードでの使用

Amazon EMR では、Spark アプリケーションを次の 2 つのモードで実行できます。

- インタラクティブ
- バッチ

コンソールまたは を使用して長時間実行されるクラスタを起動する場合 AWS CLI、SSH を使用して Hadoop ユーザーとしてマスターノードに接続し、Spark シェルを使用して Spark アプリケーションをインタラクティブに開発および実行できます。Spark をインタラクティブに使用すると、Spark アプリケーションのプロトタイプ作成やテストをバッチ環境より簡単に行うことができます。Spark アプリケーションをインタラクティブモードで正常に変更したら、そのアプリケーションの JAR または Python プログラムを、Amazon S3 上のクラスタのマスターノードに対してローカルにあるファイルシステムに配置できます。その後、アプリケーションをバッチワークフローとして送信できます。

バッチモードで、Spark スクリプトを Amazon S3 またはローカルマスターノードファイルシステムにアップロードした後、作業をクラスターにステップとして送信します。Spark ステップは、長時間稼働クラスターまたは一時的なクラスターに送信できます。

## Spark がインストールされたクラスターの作成

コンソールを使用し、Spark がインストールされた状態でクラスターを起動するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [クラスターを作成] を選択します。
3. [Software Configuration] (ソフトウェア設定) で、必要な AMI リリースバージョンを選択します。
4. [Applications to be installed] で、リストから [Spark] を選択してから、[Configure and add] を選択します。
5. 必要に応じて引数を追加して Spark 設定を変更します。詳細については、「[Spark の設定](#)」を参照してください。[Add] (追加) を選択します。
6. 必要に応じて他のオプションを選択し、[Create cluster (クラスターの作成)] を選択します。

次の例は、Java を使用した Spark のあるクラスターを作成する方法を示します。

```
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
SupportedProductConfig sparkConfig = new SupportedProductConfig()
    .withName("Spark");

RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("Spark Cluster")
    .withAmiVersion("3.11.0")
    .withNewSupportedProducts(sparkConfig)
    .withInstances(new JobFlowInstancesConfig()
        .withEc2KeyName("myKeyName")
        .withInstanceCount(1)
        .withKeepJobFlowAliveWhenNoSteps(true)
        .withMasterInstanceType("m3.xlarge")
        .withSlaveInstanceType("m3.xlarge")
    );
RunJobFlowResult result = emr.runJobFlow(request);
```

## Spark の設定

クラスターの作成時に Spark を設定するには、[Github の awslabs/emr-bootstrap-actions/spark リポジトリにある](#)ブートストラップアクションを実行します。ブートストラップアクションが受け入れる引数については、そのリポジトリの [README](#) を参照してください。ブートストラップアクションは \$SPARK\_CONF\_DIR/spark-defaults.conf ファイルのプロパティを設定します。設定の詳細については、Spark ドキュメントで、Spark 設定トピックを参照してください。次の URL の「latest」を、インストールする Spark のバージョン番号 (たとえば、2.2.0 <http://spark.apache.org/docs/latest/configuration.html>) に置き換えることができます。

各アプリケーションの送信時に、Spark を動的に設定することもできます。spark 設定ファイルを使用して、エグゼキュータのリソース割り当てを自動的に最大化する設定が使用できます。詳細については、「[Spark のデフォルト設定の上書き](#)」を参照してください。

### Spark のデフォルト設定の変更

次の例は、AWS CLIで spark.executor.memory を 2G に設定してクラスターを作成する方法を示します。

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Spark cluster" --ami-version 3.11.0 \  
--applications Name=Spark, Args=[-d,spark.executor.memory=2G] --ec2-attributes \  
KeyName=myKey \  
--instance-type m3.xlarge --instance-count 3 --use-default-roles
```

### Spark への作業の送信

作業をクラスターに送信するには、ステップを使用して EMR クラスターで spark-submit スクリプトを実行します。の addJobFlowStepsメソッドを使用してステップを追加します [AmazonElasticMapReduceClient](#)。

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
```

```
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
StepFactory stepFactory = new StepFactory();
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

StepConfig sparkStep = new StepConfig()
    .withName("Spark Step")
    .withActionOnFailure("CONTINUE")
    .withHadoopJarStep(stepFactory.newScriptRunnerStep("/home/hadoop/spark/bin/spark-submit", "--class", "org.apache.spark.examples.SparkPi", "/home/hadoop/spark/lib/spark-examples-1.3.1-hadoop2.4.0.jar", "10"));

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

## Spark のデフォルト設定の上書き

Spark のデフォルト設定値をアプリケーションごとに上書きすることができます。これは、ステップを使用してアプリケーションを送信するとき (原則としてオプションが `spark-submit` に渡されます) に行うことができます。たとえば、`spark.executor.memory` を変更することで、実行プログラムのプロセスに割り当てられたメモリを変更することができます。次のような引数を含む `--executor-memory` スイッチを指定できます。

```
/home/hadoop/spark/bin/spark-submit --executor-memory 1g --class
org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/spark-examples*.jar 10
```

同様に、`--executor-cores` と `--driver-memory` を調整できます。ステップでは、次の引数をステップに指定できます。

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/
spark-examples*.jar 10
```

さらに、`--conf` オプションを使用して、組み込みスイッチがない設定を調整することもできます。調整可能な他の設定の詳細については、Apache Spark ドキュメントの「[Spark プロパティの動的なロード](#)」を参照してください。

## Amazon EMR の以前の AMI バージョンとの S3DistCp ユーティリティの違い

### Amazon EMR でサポートされている S3DistCp バージョン

Amazon EMR AMI リリースでは、次の S3DistCp バージョンがサポートされています。1.0.7 以降の S3DistCp バージョンは、クラスターで直接確認できます。最新の機能を使用するには、`/home/hadoop/lib` の JAR を使用します。

Version	説明	リリース日
1.0.8	<code>--appendToLastFile</code> 、 <code>--requirePreviousManifest</code> 、および <code>--storageClass</code> オプションを追加します。	2014 年 1 月 3 日
1.0.7	<code>--s3ServerSideEncryption</code> オプションの追加。	2013 年 5 月 2 日
1.0.6	<code>--s3Endpoint</code> オプションの追加。	2012 年 8 月 6 日
1.0.5	実行する S3DistCp のバージョンを指定する機能が改善されました。	2012 年 6 月 27 日
1.0.4	<code>--deleteOnSuccess</code> オプションの強化。	2012 年 6 月 19 日
1.0.3	<code>--numberOfFiles</code> および <code>--startingIndex</code> オプションの追加。	2012 年 6 月 12 日
1.0.2	グループ使用時のファイル名前付け機能の強化。	2012 年 6 月 6 日
1.0.1	S3DistCp の初回リリース。	2012 年 1 月 19 日

## S3DistCp copy ステップをクラスターに追加する

実行中のクラスターに S3DistCp copy ステップを追加するには、次のコマンドを入力します。 `j-3GYXXXXXX9I0K` をクラスター ID に置き換え、 `mybucket` を Amazon S3 バケット名に置き換えます。

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr add-steps --cluster-id j-3GYXXXXXX9I0K \  
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar, \  
\   
Args=["--s3Endpoint,s3-eu-west-1.amazonaws.com",\  
"--src,s3://mybucket/logs/j-3GYXXXXXX9I0J/node/",\  
"--dest,hdfs:///output",\  
"--srcPattern,.*[a-zA-Z,]+"]
```

## Example Amazon CloudFront ログを HDFS にロードする

この例では、実行中のクラスターにステップを追加して、Amazon CloudFront ログを HDFS にロードします。このプロセスでは、圧縮形式を Gzip (デフォルト) から LZO に変更します。CloudFront。LZO を使用して圧縮されたデータは解凍時に複数のマップに分割できるので便利です。Gzip のように、圧縮が完了するのを待つ必要はありません。これにより、Amazon EMR を使用してデータを分析するときのパフォーマンスが改善します。また、この例では、`--groupBy` オプションで指定されている正規表現を使用して、指定された時間のすべてのログを 1 つのファイルにまとめることによっても、パフォーマンスが向上しています。Amazon EMR クラスターは、小さな Gzip 圧縮ファイルを多数処理するよりも、大きな数個の LZO 圧縮ファイルを処理する方が効率的に動作します。LZO ファイルを分割するには、それらのインデックスを作成して、hadoop-lzo サードパーティー製のライブラリを使用する必要があります。

Amazon CloudFront ログを HDFS にロードするには、次のコマンドを入力します。 `j-3GYXXXXXX9I0K` をクラスター ID に置き換え、 `mybucket` を Amazon S3 バケット名に置き換えます。



### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr add-steps --cluster-id j-3GYXXXXXX9I0K \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--src,s3://mybucket/cf","--dest,hdfs:///local",\
"--groupBy,. *XABCD12345678.([0-9]+-[0-9]+-[0-9]+-[0-9]+). *",\
"--targetSize,128",
"--outputCodec,lzo","--deleteOnSuccess"]
```

上記の例が次の CloudFront ログファイルで実行されている場合を考えてみましょう。

```
s3://DOC-EXAMPLE-BUCKET1/cf/XABCD12345678.2012-02-23-01.HLUS3JKx.gz
s3://DOC-EXAMPLE-BUCKET1/cf/XABCD12345678.2012-02-23-01.I9CNAZrg.gz
s3://DOC-EXAMPLE-BUCKET1/cf/XABCD12345678.2012-02-23-02.YRRwERSA.gz
s3://DOC-EXAMPLE-BUCKET1/cf/XABCD12345678.2012-02-23-02.dshVLXFE.gz
s3://DOC-EXAMPLE-BUCKET1/cf/XABCD12345678.2012-02-23-02.LpLfuShd.gz
```

S3DistCp は、ファイルを次の 2 つのファイルにコピー、連結、圧縮します。ファイル名は正規表現による一致によって決まります。

```
hdfs:///local/2012-02-23-01.lzo
hdfs:///local/2012-02-23-02.lzo
```

## 最新情報

このページでは、Amazon EMR 7.x、6.x、5.x の最新リリースで提供する変更点と機能について説明します。

これらのリリースノートは、[Amazon EMR 7.1.0](#)、[Amazon EMR 6.15.0](#)、および [Amazon EMR 5.36.2](#) ページでも、各リリースのアプリケーションバージョン、コンポーネントバージョン、および使用可能な設定分類とともに利用できます。

- これ以前のリリースのリリースノートについては、「[Amazon EMR リリースノートのアーカイブ](#)」を参照してください。
- 新しい Amazon EMR リリースが利用可能になったときに更新を取得するには、[Amazon EMR リリースノートの RSS フィード](#)にサブスクライブします。

### Note

Amazon EMR のそれ以降のリリースでは、AWS 署名バージョン 4 (SigV4) を使用して Amazon S3 へのリクエストを認証します。SigV4 をサポートする Amazon EMR リリースを使用することをお勧めします。これにより、新しい S3 バケットにアクセスでき、ワークロードの中断を回避できます。SigV4 をサポートする Amazon EMR リリースの詳細とリストについては、「[Amazon EMR と AWS 署名バージョン 4](#)」を参照してください。

## Amazon EMR 7.1.0 (7.x シリーズの最新リリース)

新しい Amazon EMR リリースは、特定のリージョンで最初に利用可能になり、その後数日で、他のリージョンでも利用できるようになります。この期間中、お客様のリージョンで最新のリリースバージョンが利用できない場合があります。

次のリリースノートには、Amazon EMR リリース 7.1.0 に関する情報が含まれています。7.1.0 からの変更が含まれています。

### 新機能

- アプリケーションのアップグレード - Amazon EMR 7.1.0 application upgrades include Livy 0.8.0, Trino 435, and ZooKeeper 3.9.1。

- [異常のあるノード交換](#) — Amazon EMR 7.1.0 以降では、異常のあるノード交換がデフォルトで有効になっているため、Amazon EMR は異常のあるノードを正常に置き換えます。Amazon EMR リリース 7.0.0 以前の既存のワークフローに影響を与えないように、クラスターで終了保護を有効にした場合、異常なノード交換は無効になります。
- [CloudWatch エージェント](#) – Amazon EMR 設定 API を使用して、追加のシステムメトリクスを使用するように CloudWatch エージェントを設定し、アプリケーションメトリクスを追加し、メトリクスの送信先を変更します。

## 既知の問題

- Python 3.11 は EMR Studio ではサポートされていません。

## 変更点、機能強化、解決した問題

- Amazon EMR 7.1.0 はデフォルトで Python 3.9 をサポートしていますが、Amazon EMR 7.1.0 の Livy 0.8.0 と Spark は Python 3.11 をサポートしています。
- このリリースでは、Python バージョン 3.11 PySpark でを使用する場合に、各行を 1 つずつ実行する必要があるという問題が修正されています。
- Zeppelin のアップグレード – Amazon EMR 7.1.0 には、Zeppelin の AWS SDK for Java v2 へのアップグレードが含まれています。このアップグレードにより、Zeppelin S3 ノートブックはカスタム暗号化マテリアルプロバイダーを受け入れることができます。AWS SDK for Java v2 は EncryptionMaterialsProvider インターフェイスを削除します。Amazon EMR 7.1.0 にアップグレードする場合、カスタム暗号化を使用する場合は、[キーリングインターフェイス](#)を実装する必要があります。Keyring インターフェイスの実装方法の例については、[KmsKeyring 「.java」](#)を参照してください。
- Amazon EMR リリース 7.1.0 にアップグレードする場合は、ローカルディスク暗号化用のカスタムキープロバイダーを変更して、AES/GCM/ の AES アルゴリズムを使用してキーを生成します NoPadding。アルゴリズムを更新しない場合、クラスターの作成が失敗してエラーが発生する可能性があります Local disk encryption failed on master instance (i-123456789) due to internal error.。カスタムキープロバイダーの作成の詳細については、[「カスタムキープロバイダーの作成」](#)を参照してください。
- Amazon EMR 7.1.0 は、開いているファイルハンドルを持つファイルのログ切り捨てロジックを改善することで、ディスク容量が少ない条件下でのノードの回復性を向上させます。

- このリリースでは、ノードの再起動時に Amazon EMR デモンの読み取り/書き込みファイルを使用して、エンコーディングとデコードのロジックを強化し、データ破損やノード障害のリスクを最小限に抑えます。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2023年3月2024日219.0	6.1.77-99.164.amzn2023	2024年5月8日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
--------------------------------------	-------------------------	-------	--------------

米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、カナダ西部 (カルガリー)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

## Amazon EMR 6.15.0 (6.x シリーズの最新リリース)

新しい Amazon EMR リリースは、特定のリージョンで最初に利用可能になり、その後数日で、他のリージョンでも利用できるようになります。この期間中、お客様のリージョンで最新のリリースバージョンが利用できない場合があります。

次のリリースノートには、Amazon EMR リリース 6.15.0 に関する情報が含まれています。6.14.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.15.0 変更ログ](#)」を参照してください。

### 新機能

- アプリケーションのアップグレード - Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426.
- [EC2 で実行される EMR クラスターの起動時間の短縮](#) - EC2 クラスターでの Amazon EMR の起動が最大 35% 速くなりました。この改善により、お客様の大多数はクラスターを 5 分以内に起動できます。

- [CodeWhisperer for EMR Studio](#) – Amazon CodeWhisperer EMR Studio で Amazon を使用して、コードを記述する際にリアルタイムのレコメンデーションを取得できるようになりました JupyterLab。CodeWhisperer は、コメントの入力、単一行のコードの完了、レコメンデーションの作成 line-by-line、および完全な形式の関数の生成を行うことができます。
- [Flink によるジョブの再起動時間の短縮](#) - Amazon EMR 6.15.0 以降では、Apache Flink にいくつかの新しいメカニズムが用意され、タスクの回復またはスケーリング操作中のジョブの再起動時間を短縮できます。これにより、実行グラフの回復と再開の速度を最適化してジョブの安定性を高めることができます。
- [オープンテーブル形式のテーブルレベルおよびきめ細かなアクセスコントロール](#) – Amazon EMR 6.15.0 以降では、AWS Glue Data Catalog のデータにアクセスする EC2 クラスターで Amazon EMR で Spark ジョブを実行するときに、AWS Lake Formation を使用して Hudi、Iceberg、または Delta Lake ベースのテーブルにテーブル、行、列、およびセルレベルのアクセス許可を適用できます。
- Hadoop のアップグレード - Amazon EMR 6.15.0 には、Apache Hadoop のバージョン 3.3.6 へのアップグレードが含まれています。Hadoop 3.3.6 は、2023 年 6 月にパッチによってリリースされた Amazon EMR 6.15 のデプロイ時の最新バージョンでした。Amazon EMR の以前のリリース (6.9.0 から 6.14.x) では Hadoop 3.3.3 が使用されていました。

このアップグレードには数百の改善と修正が含まれており、再構成可能なデータノードパラメータ、すべてのライブデータノードで一括再構成操作を開始する DFSAdmin オプション、シーク負荷の高いリーダーが複数の読み取り範囲を指定できるベクトル API などの機能が含まれています。Hadoop 3.3.6 では、HBase が他のストレージシステム実装でも実行できるように、HDFS API と先行書き込みログ (WAL) のセマンティクスのサポートも追加されています。詳細については、「Apache Hadoop ドキュメント」のバージョン [3.3.4](#)、[3.3.5](#)、[3.3.6](#) の変更ログを参照してください。

- AWS SDK for Java バージョン 2 のサポート - Amazon EMR 6.15.0 アプリケーションは、アプリケーションが v2 をサポートしている場合は AWS SDK for Java バージョン 1.12.569 または [2.20.160](#) を使用できます。 <https://github.com/aws/aws-sdk-java/tree/1.12.569> AWS SDK for Java 2.x は、バージョン 1.x のコードベースを大幅に書き換えたものです。今回のバージョンは Java 8 以降をベースに構築されており、これまで要望の声が高かった機能が複数追加されています。このバージョンでは、ノンブロッキング I/O がサポートされており、ランタイム時にさまざまな HTTP 実装をプラグインすることができます。SDK for Java v1 から v2 への移行ガイドを含めた詳細については、[AWS SDK for Java バージョン 2](#) ガイドを参照してください。

## 変更点、機能強化、解決した問題

- 高可用性 EMR クラスターを改善するため、このリリースでは、IPv6 エンドポイントを使用するローカルホスト上の Amazon EMR デモンへの接続が可能になりました。
- このリリースでは、TLS 1.2 が高可用性クラスターのすべてのプライマリノードで ZooKeeper プロビジョニングされた と通信できるようになります。
- このリリースでは、プライマリノードで管理されるトランザクションログファイルの管理 ZooKeeper が改善され、ログファイルが範囲外になり、クラスターオペレーションが中断されるシナリオが最小限に抑えられます。
- このリリースにより、高可用性 EMR クラスターのノード内通信の回復力が高まりました。この改善により、ブートストラップアクションが失敗したり、クラスターの起動に失敗したりする可能性が低くなります。
- Amazon EMR 6.15.0 の Tez では、入カスプリットを Tez グループ分割で非同期的に開くように指定できる設定が導入されています。これにより、1 つの Tez グループ分割で多数の入力分割がある場合に、読み取りクエリのパフォーマンスが向上します。詳細については、「[Tez 非同期スプリットオープン](#)」を参照してください。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>ン)、欧州(フランクフルト)、欧州(チューリッヒ)、欧州(アイルランド)、欧州(ロンドン)、欧州(パリ)、アジアパシフィック(香港)、アジアパシフィック(ムンバイ)、アジアパシフィック(ハイデラバード)、アジアパシフィック(東京)、アジアパシフィック(ソウル)、アジアパシフィック(大阪)、アジアパシフィック(シンガポール)、アジアパシフィック(シドニー)、アジアパシフィック(ジャカルタ)、アジアパシフィック(メルボルン)、アフリカ(ケープタウン)、南米(サンパウロ)、中東(バーレーン)、中東(アラブ首長国連邦)、カナダ(中部)、イスラエル(テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国(北京)、中</p>

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 13 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

## Amazon EMR 5.36.2 (5.x シリーズの最新リリース)

新しい Amazon EMR リリースは、特定のリージョンで最初に利用可能になり、その後数日で、他のリージョンでも利用できるようになります。この期間中、お客様のリージョンで最新のリリースバージョンが利用できない場合があります。

次のリリースノートには、Amazon EMR リリース 5.36.2 に関する情報が含まれています。5.36.1 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

### 変更点、機能強化、解決した問題

- このリリースでは、Amazon EMR がクラスターの HDFS レプリケーション係数設定を下回るコアノードをスケールダウンしないように、クラスターのスケールダウンロジックが改善されています。この改善により、データの冗長性要件が満たされ、スケールアップ操作が停止する可能性が軽減されます。
- このリリースでは、Presto または Trino を実行する のクラスタースケールアップワークフローに新しい再試行メカニズムが追加されました。この改善により、1 回のサイズ変更操作が失敗するため、クラスターのサイズ変更が無期限に実行されるリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。



- Amazon EMR がコアノードを正常に廃止し、完全に廃止される前に異常になる間にクラスターのスケールダウンオペレーションが停止する問題を修正しました。
- Amazon EMR が 1 つのノードを再起動するときに、複数のプライマリノードを持つ高可用性クラスター内のノードの安定性が向上します。
- Amazon EC2 で実行されている Amazon EMR でログ管理を最適化します。その結果、クラスターログのストレージコストがわずかに削減される可能性があります。
- プライマリノードで管理される ZooKeeper トランザクションログファイルの管理を改善し、ログファイルが範囲外になり、クラスターオペレーションが中断されるシナリオを最小限に抑えます。
- Yarn と通信できないために、複数のプライマリノードを持つ高可用性クラスターが失敗する可能性があるまれなバグを修正しました ResourceManager。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
2.0.2024 503.0	4.14.343	2024 年 xxxxxx 月	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバ

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
--------------------------------------	-------------------------	-------	--------------

イ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

## Amazon EMR と AWS 署名バージョン 4

Amazon EMR リリースでは、AWS 署名バージョン 4 (SigV4) を使用して Amazon S3 へのリクエストを認証します。2020 年 6 月 24 日以降に Amazon S3 で作成されたバケットは、Signature Version 2 (SigV2) によって署名されたリクエストをサポートしていません。2020 年 6 月 24 日以前に作成されたバケットは、引き続き SigV2 をサポートします。SigV4 をサポートする Amazon EMR リリースに移行することをお勧めします。これにより、新しい S3 バケットにアクセスでき、ワークロードの中断を回避できます。

Apache Spark、Apache Hive、Presto など、Amazon EMR に含まれるアプリケーションを使用している場合、SigV4 を使用するようにアプリケーションコードを変更する必要はありません。Amazon EMR に含まれていないカスタムアプリケーションを使用している場合は、SigV4 を使用するように

コードを更新する必要があるかもしれません。詳細については、「Amazon S3 ユーザーガイド」の「[Signature バージョン 2 から Signature バージョン 4 への移行](#)」を参照してください。

以下の Amazon EMR リリースは、SigV4 をサポートしています (emr-4.7.4, emr-4.8.5, emr-4.9.6, emr-4.10.1, emr-5.1.1, emr-5.2.3, emr-5.3.2, emr-5.4.1, emr-5.5.4, emr-5.6.1, emr-5.7.1, emr-5.8.3, emr-5.9.1, emr-5.10.1, emr-5.11.4, emr-5.12.3, emr-5.13.1, emr-5.14.2, emr-5.15.1, emr-5.16.1, emr-5.17.2, emr-5.18.1, emr-5.19.1, emr-5.20.1, emr-5.21.2, and emr-5.22.0 and higher)。6.x と 7.x のすべてのリリースが SigV4 をサポートしています。

## CVE-2021-44228 を軽減するためのアプローチ

### Note

Amazon EMR リリース 6.9.0 以降では、Amazon EMR によってインストールされ、Log4j ライブラリを使用するすべてのコンポーネントは Log4j バージョン 2.17.1 以降を使用します。

EC2 で実行されている Amazon EMR

[CVE-2021-44228](#) で説明されている問題は、信頼できないソースからの入力を処理する場合の Apache Log4j コアバージョン 2.0.0 から 2.14.1 までに関係するものです。5.34.0 までの Amazon EMR 5.x リリースと Amazon EMR 6.5.0 までの EMR 6.x リリースで起動された Amazon EMR クラスターには、これらのバージョンの Apache Log4j を使用する、Apache Hive、Flink、HUDI、Presto、Trino などのオープンソースフレームワークが含まれています。ただし、多くのお客様は、Amazon EMR クラスターにインストールされているオープンソースフレームワークを使用して、信頼できないソースからの入力を処理し、ログ記録しています。

以下のセクションで説明するように、「Log4j CVE-2021-44228 用 Amazon EMR ブートストラップアクションソリューション」を適用することをお勧めします。このソリューションは CVE-2021-45046 にも対応しています。

### Note

Amazon EMR のブートストラップアクションスクリプトは 2022 年 9 月 7 日に更新され、Oozie の段階的なバグ修正と改善が含まれています。Oozie を使用する場合は、以下のセクションで説明する更新された Amazon EMR ブートストラップアクションソリューションを適用する必要があります。

## Amazon EMR on EKS

[Amazon EMR on EKS](#) をデフォルト設定で使用する場合、CVE-2021-44228 で説明されている問題の影響を受けず、「[Log4j CVE-2021-44228 および CVE-2021-45046 用 Amazon EMR ブートストラップアクションソリューション](#)」セクションで説明されているソリューションを適用する必要もありません。Amazon EMR on EKS の場合、Spark の Amazon EMR ランタイムは Apache Log4j バージョン 1.2.17 を使用します。Amazon EMR on EKS を使用するときは、log4j.appender コンポーネントのデフォルト設定を log に変更しないでください。

## Log4j CVE-2021-44228 および CVE-2021-45046 用 Amazon EMR ブートストラップアクションソリューション

このソリューションでは、Amazon EMR クラスターに適用する必要がある Amazon EMR ブートストラップアクションを提供します。以下に、Amazon EMR のリリースごとに、ブートストラップアクションスクリプトへのリンクが示してあります。このブートストラップアクションを適用するには、次の手順を完了する必要があります。

1. Amazon EMR リリースに対応するスクリプトを、AWS アカウントのローカル S3 バケットにコピーします。お使いの Amazon EMR リリース固有のブートストラップスクリプトを使用していることを確認してください。
2. [EMR ドキュメント](#)に記載されている手順に従って、EMR クラスターのブートストラップアクションを設定し、S3 バケットにコピーされたスクリプトを実行します。EMR クラスターに他のブートストラップアクションが設定されている場合は、このスクリプトが、最初に実行するブートストラップアクションスクリプトとして設定されていることを確認してください。
3. 既存の EMR クラスターを終了し、ブートストラップアクションスクリプトを使用して新しいクラスターを起動します。テスト環境でブートストラップスクリプトを AWS テストし、本番環境に適用する前にアプリケーションを検証することをお勧めします。EMR マイナーリリースの最新リビジョン (6.3.0 など) を使用していない場合は、最新リビジョン (6.3.1 など) を使用し、前述のソリューションを適用する必要があります。

### CVE-2021-44228 および CVE-2021-45046 - Amazon EMR リリース用ブートストラップスクリプト

Amazon EMR リリース番号	スクリプトの場所	スクリプトリリース日
6.5.0	s3://elasticmapreduce/ bootstrap-actions/	2022 年 3 月 24 日

Amazon EMR リリース番号	スクリプトの場所	スクリプトリリース日
	<code>log4j/patch-log4j-emr-6.5.0-v2.sh</code>	
6.4.0	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j-emr-6.4.0-v2.sh</code>	2022 年 3 月 24 日
6.3.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j-emr-6.3.1-v2.sh</code>	2022 年 3 月 24 日
6.2.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j-emr-6.2.1-v2.sh</code>	2022 年 3 月 24 日
6.1.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j-emr-6.1.1-v2.sh</code>	2021 年 12 月 14 日
6.0.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j-emr-6.0.1-v2.sh</code>	2021 年 12 月 14 日
5.34.0	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j-emr-5.34.0-v2.sh</code>	2021 年 12 月 12 日

Amazon EMR リリース番号	スクリプトの場所	スクリプトリリース日
5.33.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.33.1-v2.sh</code>	2021 年 12 月 12 日
5.32.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.32.1-v2.sh</code>	2021 年 12 月 13 日
5.31.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.31.1-v2.sh</code>	2021 年 12 月 13 日
5.30.2	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.30.2-v2.sh</code>	2021 年 12 月 14 日
5.29.0	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.29.0-v2.sh</code>	2021 年 12 月 14 日
5.28.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.28.1-v2.sh</code>	2021 年 12 月 15 日
5.27.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.27.1-v2.sh</code>	2021 年 12 月 15 日

Amazon EMR リリース番号	スクリプトの場所	スクリプトリリース日
5.26.0	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.26.0-v2.sh</code>	2021 年 12 月 15 日
5.25.0	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.25.0-v2.sh</code>	2021 年 12 月 15 日
5.24.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.24.1-v2.sh</code>	2021 年 12 月 15 日
5.23.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.23.1-v2.sh</code>	2021 年 12 月 15 日
5.22.0	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.22.0-v2.sh</code>	2021 年 12 月 15 日
5.21.2	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.21.2-v2.sh</code>	2021 年 12 月 15 日
5.20.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.20.1-v2.sh</code>	2021 年 12 月 15 日

Amazon EMR リリース番号	スクリプトの場所	スクリプトリリース日
5.19.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.19.1-v2.sh</code>	2021 年 12 月 15 日
5.18.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.18.1-v2.sh</code>	2021 年 12 月 15 日
5.17.2	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.17.2-v2.sh</code>	2021 年 12 月 15 日
5.16.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.16.1-v2.sh</code>	2021 年 12 月 15 日
5.15.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.15.1-v2.sh</code>	2021 年 12 月 15 日
5.14.2	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.14.2-v2.sh</code>	2021 年 12 月 15 日
5.13.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.13.1-v2.sh</code>	2021 年 12 月 15 日



Amazon EMR リリース番号	スクリプトの場所	スクリプトリリース日
5.12.3	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.12.3-v2.sh</code>	2021 年 12 月 15 日
5.11.4	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.11.4-v2.sh</code>	2021 年 12 月 15 日
5.10.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.10.1-v2.sh</code>	2021 年 12 月 15 日
5.9.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.9.1-v2.sh</code>	2021 年 12 月 15 日
5.8.3	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.8.3-v2.sh</code>	2021 年 12 月 15 日
5.7.1	<code>s3://elasticmapreduce/ bootstrap-actions/ log4j/patch-log4j- emr-5.7.1-v2.sh</code>	2021 年 12 月 15 日

EMR リリースバージョン	2021 年 12 月現在の最新リビジョン
6.3.0	6.3.1

EMR リリースバージョン	2021 年 12 月現在の最新リビジョン
6.2.0	6.2.1
6.1.0	6.1.1
6.0.0	6.0.1
5.33.0	5.33.1
5.32.0	5.32.1
5.31.0	5.31.1
5.30.0 または 5.30.1	5.30.2
5.28.0	5.28.1
5.27.0	5.27.1
5.24.0	5.24.1
5.23.0	5.23.1
5.21.0 または 5.21.1	5.21.2
5.20.0	5.20.1
5.19.0	5.19.1
5.18.0	5.18.1
5.17.0 または 5.17.1	5.17.2
5.16.0	5.16.1
5.15.0	5.15.1
5.14.0 または 5.14.1	5.14.2
5.13.0	5.13.1

EMR リリースバージョン	2021 年 12 月現在の最新リビジョン
5.12.0、5.12.1、5.12.2	5.12.3
5.11.0、5.11.1、5.11.2、5.11.3	5.11.4
5.9.0	5.9.1
5.8.0、5.8.1、5.8.2	5.8.3
5.7.0	5.7.1

## よくある質問

- EMR 5 より古い EMR リリースは CVE-2021-44228 の影響を受けますか？

いいえ。EMR リリース 5 より前の EMR リリースでは、2.0 より古い Log4j バージョンが使用されています。

- このソリューションは CVE-2021-45046 に対応していますか？

はい。このソリューションは CVE-2021-45046 にも対応しています。

- このソリューションは、私が自分の EMR クラスターにインストールしたカスタムアプリケーションを処理しますか？

ブートストラップスクリプトは、EMR によってインストールされた JAR ファイルのみを更新します。ブートストラップアクション、クラスターに送信されたステップ、カスタム Amazon Linux AMI の使用、またはその他のメカニズムを通じて、カスタムアプリケーションと JAR ファイルを EMR クラスターにインストールして実行する場合は、アプリケーションベンダーと協力して、カスタムアプリケーションが CVE-2021-44228 の影響を受けるかどうかを判断し、適切な解決策を決定してください。

- EMR on EKS を使用して、[カスタマイズされた docker イメージ](#)をどのように処理すればよいですか？

[カスタマイズされた docker イメージ](#)を使用して Amazon EMR on EKS にカスタムアプリケーションを追加するか、カスタムアプリケーションファイルを使用して Amazon EMR on EKS にジョブを送信する場合は、アプリケーションベンダーと協力して、カスタムアプリケーションが CVE-2021-44228 の影響を受けるかどうかを判断し、適切な解決策を決定してください。

- ブートストラップスクリプトは、CVE-2021-44228 および CVE-2021-45046 で説明されている問題を軽減するためにどのように機能しますか？

ブートストラップスクリプトは、新しい命令セットを追加して EMR 起動命令を更新します。これらの新しい手順では、EMR によってインストールされたすべてのオープンソースフレームワークによって Log4j を介して使用される JndiLookup クラスファイルを削除します。これは、Log4j の問題に対処するための [Apache の勧告](#) に従うものです。

- Log4j バージョン 2.17.1 以降を使用する EMR の更新はありますか？

リリース 5.34 までの EMR 5 リリースとリリース 6.5 までの EMR 6 リリースでは、Log4j の最新バージョンと互換性のない古いバージョンのオープンソースフレームワークが使用されています。これらのリリースを引き続き使用する場合は、CVE で説明されている問題を軽減するためにブートストラップアクションを適用することをお勧めします。EMR 5 リリース 5.34 と EMR 6 リリース 6.5 以降、Log4j 1.x と Log4j 2.x を使用するアプリケーションは、それぞれ Log4j 1.2.17 (またはそれ以上) と Log4j 2.17.1 (またはそれ以上) を使用するようアップグレードされます。CVE の問題を軽減するために、前述のブートストラップアクションを使用する必要はなくなります。

- EMR リリースは CVE-2021-45105 の影響を受けますか？

EMR のデフォルト設定で Amazon EMR によってインストールされたアプリケーションは、CVE-2021-45105 の影響を受けません。Amazon EMR によってインストールされたアプリケーションのうち、[コンテキスト検索](#)で Apache Log4j を使用するのは Apache Hive のみであり、不適切な入力データを処理するようなデフォルト以外のパターンレイアウトは使用していません。

- Amazon EMR は、以下の CVE 開示のいずれかによって影響を受けますか？

次の表には、Log4j に関連する CVE のリストと、各 CVE が Amazon EMR に影響するかどうか記載されています。この表の情報は、Amazon EMR がデフォルト設定を使用してアプリケーションをインストールする場合にのみ適用されます。

CVE	EMR への影響	メモ
CVE-2022-23302	なし	Amazon EMR は Log4j JMSSink をセットアップしません
CVE-2022-23305	なし	Amazon EMR は Log4j JDBCAppender をセットアップしません

CVE	EMR への影響	メモ
CVE-2022-23307	なし	Amazon EMR は Log4j Chainsaw をセットアップしません
CVE-2020-9493	なし	Amazon EMR は Log4j Chainsaw をセットアップしません
CVE-2021-44832	なし	Amazon EMR は JNDI 接続文字列を使用して Log4j JDBCAppender をセットアップしません
CVE-2021-4104	なし	Amazon EMR は Log4j JMSAppender を使用しません
CVE-2020-9488	なし	Amazon EMR によってインストールされるアプリケーションは Log4j SMTPAppender を使用しません
CVE-2019-17571	なし	Amazon EMR はクラスターへのパブリックアクセスをブロックし、起動しません SocketServer

CVE	EMR への影響	メモ
CVE-2019-17531	なし	最新の Amazon EMR リリースバージョンにアップグレードすることをお勧めします。Amazon EMR 5.33.0 以降は jackson-databind 2.6.7.4 以降を使用し、EMR 6.1.0 以降は jackson-databind 2.10.0 以降を使用します。これらのバージョンの jackson-databind は CVE の影響を受けません。

## Amazon EMR リリースノートのアーカイブ

すべての Amazon EMR リリースのリリースノートは以下から入手できます。各リリースの包括的なリリース情報については、「[Amazon EMR 6.x リリースバージョン](#)」、「[Amazon EMR 5.x リリースバージョン](#)」、および「[Amazon EMR 4.x リリースバージョン](#)」を参照してください。

新しい Amazon EMR リリースが利用可能になったときに更新を取得するには、[Amazon EMR リリースノートの RSS フィード](#)にサブスクライブします。

### リリース 6.14.0

次のリリースノートには、Amazon EMR リリース 6.14.0 に関する情報が含まれています。6.13.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.14.0 変更ログ](#)」を参照してください。

#### 新機能

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- Amazon EMR 6.14.0 以降で作成したクラスターについて、ap-southeast-3 アジアパシフィック (ジャカルタ) リージョンで [Amazon EMR Managed Scaling](#) を利用できるようになりました。

## 変更点、機能強化、解決した問題

- 6.14.0 リリースでは、Amazon EC2 で実行されている Amazon EMR でのログ管理が最適化されています。その結果、クラスターログのストレージコストがわずかに削減される可能性があります。
- 6.14.0 リリースでは、Amazon EBS ボリュームのサイズに大きなばらつきがあるさまざまなコアインスタンスを考慮して、スケーリングワークフローが改善されています。この改善はコアノードにのみ適用され、タスクノードのスケールダウン操作には影響しません。
- 6.14.0 リリースでは、Amazon EMR が Apache Hadoop YARN ResourceManager and HDFS NameNode などのオープンソースアプリケーションとやり取りする方法が改善されています。この改善により、クラスタースケーリングによる運用遅延のリスクが軽減され、オープンソースアプリケーションとの接続の問題が原因で発生する起動障害が軽減されます。
- 6.14.0 リリースでは、クラスター起動時のアプリケーションのインストールが最適化されます。これにより、Amazon EMR アプリケーションの特定の組み合わせのクラスター起動時間が短縮されます。
- 6.14.0 リリースでは、カスタムドメインの VPC で実行されているクラスターがコアノードまたはタスクノードの再起動に遭遇したときに、クラスターのスケールダウン操作が停止することがある問題が修正されています。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			(ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 17 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 9 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

## リリース 6.13.0

次のリリースノートには、Amazon EMR リリース 6.13.0 に関する情報が含まれています。6.12.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.13.0 変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

### 変更点、機能強化、解決した問題

- 6.13.0 リリースでは、Amazon EMR ログ管理デーモンが改善され、クラスター終了コマンドが発行されたときに、すべてのログが定期的に Amazon S3 にアップロードされるようになりました。これにより、クラスターの終了が速くなります。
- 6.13.0 リリースでは Amazon EMR のログ管理機能が強化され、すべてのログファイルを Amazon S3 に一貫してタイムリーにアップロードできるようになりました。これは特に、長時間稼働する EMR クラスターにメリットをもたらします。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

## リリース 6.12.0

次のリリースノートには、Amazon EMR リリース 6.12.0 に関する情報が含まれています。6.11.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.12.0 変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- Amazon EMR リリース 6.12.0 以降では、Apache Livy、Apache Hive through HiveServer2 (HS2)、Trino、Presto、Hue との LDAP 統合がサポートされています。6.12.0 以降を使用する EMR クラスターに Apache Spark と Apache Hadoop をインストールし、LDAP を使用するように設定することもできます。詳細については、「[Amazon EMR での認証に Active Directory または LDAP サーバーを使用する](#)」を参照してください。

### 変更点、機能強化、解決した問題

- Amazon EMR リリース 6.12.0 以降では、Java 11 ランタイム環境で Flink を実行できます。詳細については、「[Flink が Java 11 で実行されるよう設定する](#)」を参照してください。
- 6.12.0 リリースでは、Presto または Trino を実行する EMR クラスターのクラスタースケリングワークフローに新しい再試行メカニズムが追加されています。この改善により、サイズ変更操作が 1 回失敗したためにクラスターのサイズ変更が無期限に停止するリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。

- 6.12.0 リリースでは、正常に廃止されているコアノードが完全に廃止される前に何らかの理由で異常が発生すると、クラスターのスケールダウン操作が停止することがある問題が修正されています。
- 6.12.0 リリースでは、クラスターのスケールダウンロジックが改善され、クラスターの HDFS レプリケーション係数設定を下回るコアノードのスケールダウンが試みられることがなくなりました。これはデータの冗長性要件に合致し、スケールダウン操作が停止する可能性が低くなります。
- 6.12.0 リリースでは、インスタンスの状態変化を記録する速度が向上し、Amazon EMR のヘルスマニタリングサービスのパフォーマンスと効率が向上します。この改善により、複数のカスタムクライアントツールやサードパーティーアプリケーションを実行しているクラスターノードのパフォーマンスが低下する可能性が低くなります。
- 6.12.0 リリースでは、Amazon EMR のクラスター上のログ管理デーモンのパフォーマンスが向上しています。その結果、同時実行性の高いステップを実行する EMR クラスターでは、パフォーマンスが低下する可能性が低くなります。
- Amazon EMR リリース 6.12.0 では、ログ管理デーモンがアップグレードされ、ローカルインスタンスストレージ上のオープンファイルハンドルでアクティブに使用されているすべてのログと、関連するプロセスが識別されるようになりました。このアップグレードにより、ログが Amazon S3 にアーカイブされた後に Amazon EMR がファイルを適切に削除し、ストレージスペースを再利用できるようになります。
- 6.12.0 リリースには、ローカルクラスターファイルシステム内の空で未使用のステップディレクトリを削除するログ管理デーモンの機能強化が含まれています。空のディレクトリが多すぎると、Amazon EMR デーモンのパフォーマンスが低下し、ディスクが過剰に使用される可能性があります。
- 6.12.0 リリースでは、YARN タイムラインサーバーのログのログローテーションが可能になります。これにより、特に長時間稼働するクラスターのディスク過剰使用シナリオが最小限に抑えられます。
- Amazon EMR 6.10.0 以降では、デフォルトのルートボリュームサイズが 15 GB に増えました。以前のリリースでは、デフォルトのルートボリュームサイズは 10 GB でした。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)

## リリース 6.11.1

次のリリースノートには、Amazon EMR リリース 6.11.1 に関する情報が含まれています。6.11.0 からの変更が含まれています。リリースタイムラインの詳細については、「[6.11.1 変更ログ](#)」を参照してください。

### 変更点、機能強化、解決した問題

- ノードが廃止試行と同時に追加または削除されると、ロック競合が原因でデッドロックに陥る可能性があります。その結果、Hadoop Resource Manager (YARN) が応答しなくなり、着信コンテナと現在実行中のコンテナのすべてに影響が及びます。
- このリリースには、高可用性クラスターが再起動後に障害状態から回復できるようにする変更が含まれています。
- このリリースには Hue と HBase のセキュリティ修正が含まれています。
- このリリースでは、Amazon EMR で Spark でワークロードを実行しているクラスターが `contains`、`startsWith`、`endsWith`、および `like` でメッセージが表示されずに間違った結果を受け取る可能性があるという問題が修正されています。この問題は、Amazon EMR Hive3 メタストアサーバー (HMS) のメタデータを含むパーティション化されたフィールドで式を使用する場合に発生します。
- このリリースでは、ユーザー定義関数 (UDF) がない場合の Glue 側のスロットリングの問題が修正されています。
- このリリースでは、YARN が廃止された場合に、ログプッシャーが S3 にコンテナログをプッシュする前に、ノードログ集約サービスによってコンテナログが削除される問題が修正されています。
- このリリースでは、Hadoop で Node Label が有効になっている場合のス FairShare ケジューラメトリクスの問題が修正されています。
- このリリースでは、`spark-defaults.conf` の `spark.yarn.heterogeneousExecutors.enabled` 設定にデフォルトの `true` 値を設定したときに Spark のパフォーマンスに影響する問題が修正されています。
- このリリースでは、低減タスクがシャッフルデータの読み取りに失敗する問題が修正されています。この問題により、メモリ破損エラーで Hive のクエリが失敗していました。
- このリリースでは、Presto または Trino を実行する EMR クラスターのクラスタースケールアップワークフローに新しい再試行メカニズムが追加されています。この改善により、サイズ変更操作が 1 回失敗したためにクラスターのサイズ変更が無期限に停止するリスクが軽減されます。また、クラスターのスケールアップとスケールダウンが速くなるため、クラスターの使用率も向上します。

- このリリースでは、クラスターのスケールダウンロジックが改善され、クラスターの HDFS レプリケーション係数設定を下回るコアノードのスケールダウンが試みられることがなくなりました。これはデータの冗長性要件に合致し、スケールダウン操作が停止する可能性が低くなります。
- ログ管理デーモンがアップグレードされ、ローカルインスタンスストレージ上のオープンファイルハンドルでアクティブに使用されているすべてのログと、関連するプロセスが識別されるようになりました。このアップグレードにより、ログが Amazon S3 にアーカイブされた後に Amazon EMR がファイルを適切に削除し、ストレージスペースを再利用できるようになります。
- このリリースには、ローカルクラスターファイルシステム内の空で未使用のステップディレクトリを削除するログ管理デーモンの機能強化が含まれています。空のディレクトリが多すぎると、Amazon EMR デーモンのパフォーマンスが低下し、ディスクが過剰に使用される可能性があります。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 4.14.336 223.0		2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパ

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>シフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)</p>



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)、カナダ西部 (カルガリー)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バージェン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)、イスラエル (テルアビブ)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 4.14.320 727.0		2023 年 8 月 14 日	連邦)、カナダ (中部)、イスラエル (テルアビブ)  米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

## リリース 6.11.0

次のリリースノートには、Amazon EMR リリース 6.11.0 に関する情報が含まれています。6.10.0 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR 6.11.0 は、Apache Spark 3.3.2-amzn-0、Apache Spark RAPIDS 23.02.0-amzn-0、CUDA 11.8.0、Apache Hudi 0.13.0-amzn-0、Apache Iceberg 1.2.0-amzn-0、Trino 410-amzn-0、PrestoDB 0.279-amzn-0 をサポートしています。

### 変更点、機能強化、解決した問題

- Amazon EMR 6.11.0 では、DynamoDB コネクタがバージョン 5.0.0 にアップグレードされています。バージョン 5.0.0 では `aws.sdk.java` を使用します AWS SDK for Java 2.x。以前のリリースでは AWS SDK for Java 1.x が使用されていました。このアップグレードの結果として、Amazon EMR 6.11 で DynamoDB コネクタを使用する前にコードをテストすることを強くお勧めします。
- Amazon EMR 6.11.0 の DynamoDB コネクタは、DynamoDB サービスを呼び出すとき、`dynamodb.endpoint` プロパティに指定されたリージョンの値を使用します。`dynamodb.endpoint` を使用するとき `dynamodb.region` も設定し、両方のプロパティで同じ AWS リージョンをターゲットにすることをお勧めします。`dynamodb.endpoint` を設定しない場合 `dynamodb.region`、Amazon EMR 6.11.0 の DynamoDB コネクタは無効なリージョン例外を返し、Amazon EC2 インスタンスメタデータサービス (IMDS) からの AWS リージョン 情報の照合を試みます。コネクタが IMDS からリージョンを取得できない場合は、デフォルトで米国東部 (バージニア北部) (`us-east-1`) になります。次のエラーは、`dynamodb.region` プロパティを適切に設定しない場合に発生する可能性のある無効なリージョン例外の例です。`software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` AWS SDK for Java へのアップグレードの影響を受けるクラスの詳細については、Amazon EMR - DynamoDB コネクタの GitHub リポジトリの「[1.x AWS SDK for Java から 2.x \(#175\) へのアップグレード](#)」コミットを参照してください。
- このリリースでは、列の名前変更操作後に Delta Lake を使用して Amazon S3 に Delta テーブルデータを保存すると列データが NULL になる問題が修正されています。Delta Lake のこの実験的機能の詳細については、「Delta Lake User Guide」の「[Column rename operation](#)」を参照してください。

- 6.11.0 リリースでは、複数のプライマリノードがあるクラスターからプライマリノードの 1 つを複製してエッジノードを作成する際に発生する可能性のある問題が修正されています。複製されたエッジノードは、スケールダウン操作で遅延を引き起こしたり、プライマリノードのメモリ使用率が高くなったりする可能性があります。EMR クラスターと通信するエッジノードを作成する方法の詳細については、の [aws-samples](#) リポジトリの「[エッジノードの作成](#)」を参照してください [GitHub](#)。
- 6.11.0 リリースでは、Amazon EMR が再起動後に Amazon EBS ボリュームをインスタンスに再マウントするために使用する自動化プロセスが改善されました。
- 6.11.0 リリースでは、Amazon EMR が Amazon に発行する Hadoop メトリクスに断続的なギャップが生じる問題が修正されています [CloudWatch](#)。
- 6.11.0 リリースでは、クラスターのノードの除外リストを含む YARN 構成ファイルの更新が、ディスクの過剰使用により中断される EMR クラスターの問題が修正されています。更新が不完全だと、今後のクラスターのスケールダウン操作が妨げられます。このリリースでは、クラスターが正常に動作し、スケーリング操作が期待どおりに機能することが保証されます。
- Amazon EMR 6.10.0 以降では、デフォルトのルートボリュームサイズが 15 GB に増えました。以前のリリースでは、デフォルトのルートボリュームサイズは 10 GB でした。
- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。このアプローチでは、マネージドスケーリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。

Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために `yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値が `false` に設定されています。

この修正は YARN-9608 によって発生した問題を解決するものですが、マネージドスケーリングが有効になっているクラスターでのシャッフルデータ損失が原因で Hive ジョブが失敗する可能性があります。このリリースでは、Hive ワークロードにも `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` を設定することで、このリスクを軽減しました。この設定は、Amazon EMR リリース 6.11.0 以降でのみ使用できます。

- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

**Note**

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または [ListReleaseLabels](#) API あるいは [list-release-labels](#) CLI アクションを使用するときに Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)

## リリース 6.10.0

次のリリースノートには、Amazon EMR リリース 6.10.0 に関する情報が含まれています。6.9.0 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR 6.10.0 は、Apache Spark 3.3.1、Apache Spark RAPIDS 22.12.0、CUDA 11.8.0、Apache Hudi 0.12.2-amzn-0、Apache Iceberg 1.1.0-amzn-0、Trino 403、PrestoDB 0.278.1 をサポートしています。
- Amazon EMR 6.10.0 には、Hudi テーブル内のデータへの読み取りアクセスを提供するネイティブ Trino-Hudi コネクタが含まれています。コネクタは `trino-cli --catalog hudi` で有効化でき、`trino-connector-hudi` との要件に合わせてコネクタを設定できます。Amazon EMR とのネイティブ統合により、Hudi テーブルのクエリに `trino-connector-hive` を使用する必要がなくなりました。新しいコネクタでサポートされる設定のリストについては、Trino のドキュメントの「[Hudi connector](#)」ページを参照してください。
- Amazon EMR リリース 6.10.0 以降では、Apache Zeppelin と Apache Flink の統合がサポートされています。詳細については、「[Amazon EMR の Zeppelin から Flink ジョブを操作する](#)」を参照してください。

### 既知の問題

- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。このアプローチでは、マネージドスケーリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。

Amazon EMR 6.10.0 でこの問題を回避するには、`yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値を `false` に設定します。Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために設定がデフォルトで `false` にセットされています。

## 変更点、機能強化、解決した問題

- Amazon EMR 6.10.0 では、[Apache Spark 用の Amazon Redshift の統合](#)のために `minimal-json.jar` の依存関係が削除され、必要な Spark-Redshift 関連の jar が Spark のエグゼキュータークラスパス `spark-redshift.jar`、`spark-avro.jar`、および `RedshiftJDBC.jar` に自動的に追加されます。
- 6.10.0 リリースでは、クラスター上のログ管理デーモンが改善され、EMR クラスター内の追加のログフォルダを監視できるようになりました。この改善により、ディスクの過剰使用シナリオが最小限に抑えられます。
- 6.10.0 リリースでは、クラスター上のログ管理デーモンが停止すると、自動的に再起動されます。この改善により、ディスクの過剰使用が原因でノードが異常に見えるリスクが軽減されます。
- Amazon EMR 6.10.0 は EMRFS ユーザーマッピング用のリージョンのエンドポイントをサポートしています。
- Amazon EMR 6.10.0 以降では、デフォルトのルートボリュームサイズが 15 GB に増えました。以前のリリースでは、デフォルトのルートボリュームサイズは 10 GB でした。
- 6.10.0 リリースでは、残りのすべての Spark エグゼキューターが YARN リソースマネージャーを備えた廃止予定のホスト上にあるとき Spark ジョブが停止する問題が修正されています。
- Amazon EMR 6.6.0 から 6.9.x では、動的パーティションと ORDER BY 句または SORT BY 句を使用した INSERT クエリには常に 2 つのリデューサーがあります。この問題は、OSS が変更された [HIVE-20703](#) が原因です。これにより、動的ソートパーティションの最適化がコストベースの決定下に置かれます。ワークロードで動的パーティションのソートが不要な場合は、`hive.optimize.sort.dynamic.partition.threshold` プロパティを `-1` に設定して新機能を無効にし、リデューサーの数を正しく計算することをお勧めします。この問題は、[HIVE-22269](#) の一部として OSS Hive で修正され、Amazon EMR 6.10.0 で修正されています。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

### Note

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または

[ListReleaseLabels](#) API あるいは [list-release-labels](#) CLI アクションを使用するときに Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (ミラノ)、欧州 (スペイン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (UAE)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (UAE)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202307.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (UAE)

## リリース 6.9.0

次のリリースノートには、Amazon EMR リリース 6.9.0 に関する情報が含まれています。Amazon EMR リリース 6.8.0 からの変更が含まれています。リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

### 新機能

- Amazon EMR リリース 6.9.0 は、Apache Spark RAPIDS 22.08.0、Apache Hudi 0.12.1、Apache Iceberg 0.14.1、Trino 398、Tez 0.10.2 をサポートしています。
- Amazon EMR リリース 6.9.0 には、新しいオープンソースアプリケーションである [Delta Lake 2.1.0](#) が含まれています。
- Apache Spark 用の Amazon Redshift の統合は、Amazon EMR リリース 6.9.0 以降に含まれています。以前はオープンソースツールであったこのネイティブインテグレーションは Spark コネクタと呼ばれるもので、これを使用して Apache Spark アプリケーションを構築することで、Amazon Redshift と Amazon Redshift Serverless 内のデータを読み書きできます。詳細については、「[Amazon EMR での Apache Spark 用の Amazon Redshift インテグレーションの使用](#)」を参照してください。
- Amazon EMR リリース 6.9.0 では、クラスターのスケールダウン中に Amazon S3 にログをアーカイブするためのサポートが追加されています。以前は、Amazon S3 にログファイルをアーカイブできるのはクラスター終了時のみでした。この新機能により、クラスターで生成されたログファイルは、ノードが終了した後も Amazon S3 に残ります。詳細については、「[クラスターのログ記録とデバッグを設定する](#)」を参照してください。
- 長時間実行されるクエリをサポートするため、Trino には耐障害性実行メカニズムが組み込まれました。耐障害性実行では、失敗したクエリやそのコンポーネントタスクを再試行することで、クエリの失敗を軽減します。詳細については、「[Trino での耐障害性実行](#)」を参照してください。
- Amazon EMR で Apache Flink を使用して、Apache Hive テーブル、または Iceberg、Kinesis、Kafka などの任意の Flink テーブルソースのメタデータの BATCH 処理と STREAM 処理を統合することができます。、または Amazon EMR API を使用して AWS Management Console AWS CLI、AWS Glue データカタログを Flink のメタストアとして指定できます。詳細については、「[Amazon EMR で Flink を設定する](#)」を参照してください。
- Amazon SageMaker Studio を使用して EC2 クラスター上の Amazon EMR で Apache Spark、Apache Hive、Presto クエリに AWS Identity and Access Management (IAM) ランタイムロールと AWS Lake Formation ベースのアクセスコントロールを指定できるようになりました。詳細については、「[Configure runtime roles for Amazon EMR steps](#)」を参照してください。

## 既知の問題

- Amazon EMR リリース 6.9.0 では、Apache Ranger が有効になっているクラスターでは Trino は動作しません。Ranger で Trino を使用する必要がある場合は、[AWS Support](#) にお問い合わせください。
- Amazon Redshift integration for Apache Spark を使用している場合に、time、timetz、timestamp、timestampz のいずれかにマイクロ秒の精度を Parquet 形式で設定していると、コネクタがその時間値を最も近いミリ秒値に四捨五入します。回避策として、テキストアンロード形式 unload\_s3\_format パラメータを使用してください。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、s3://bucket/table/p=a は s3://bucket/table/p=a b のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の

spark.sql.sources.fastS3PartitionDiscovery.enabled 設定を false にセットします。

- Amazon SageMaker Studio から Amazon EMR クラスターへの接続は、403 Forbidden レスポンスコードで断続的に失敗することがあります。このエラーは、クラスターでの IAM ロールの設定に 60 秒以上かかる場合に発生します。回避策として、Amazon EMR パッチをインストールして再試行を有効にし、タイムアウトを最低 300 秒に増やすことができます。以下の手順を使用して、クラスターの起動時にブートストラップアクションを適用します。

1. 次の Amazon S3 URI から、ブートストラップスクリプトと RPM ファイルをダウンロードします。

```
s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
```

```
s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
```

2. 前のステップのファイルを、所有する Amazon S3 バケットにアップロードします。バケットは、クラスターを起動する AWS リージョン 予定のと同じにある必要があります。
3. EMR クラスターを起動するときに、次のブートストラップアクションを含めません。 *bootstrap\_URI* と *RPM\_URI* を Amazon S3 の対応する URI に置き換えます。

```
--bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
```

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、Log4j2 プロパティのファイル名パターン設定が正しくないため、SecretAgent および RecordServer サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- Apache Flink はネイティブ S3 FileSystem および Hadoop FileSystem コネクタを提供します。これにより、アプリケーションはを作成し、データを Amazon S3 に FileSink 書き込むことができます。これは、次の 2 つの例外のいずれかで FileSink 失敗します。

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError: org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/
```

```
apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/
RetryPolicy;

                                at
  org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-
yarn-common-3.3.3-amzn-0.jar:?]
```

回避策として、Flink の上記の問題を修正する Amazon EMR パッチをインストールできます。クラスターを起動するときにはブートストラップアクションを適用するには、以下の手順を実行します。

1. flink-rpm を Amazon S3 バケットにダウンロードします。RPM パスは `s3://DOC-EXAMPLE-BUCKET/repms/flink/` です。
2. 次の URI を使用して Amazon S3 からブートストラップスクリプトと RPM ファイルをダウンロードします。を、クラスターを起動する AWS リージョン 予定の `regionName` に置き換えます。

```
s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-
rpms.sh
```

3. Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。Amazon EMR 6.8.0 および 6.9.0 の場合、この方法では、マネージドスケーリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性があります。

[Amazon EMR 6.10.0](#) では、`yarn-site.xml` の

`yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値を `false` に設定することでこの問題を回避できます。Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために設定がデフォルトで `false` にセットされています。

## 変更、拡張、解決した問題

- Amazon EMR リリース 6.9.0 以降では、Amazon EMR によってインストールされ、Log4j ライブラリを使用するすべてのコンポーネントは Log4j バージョン 2.17.1 以降を使用します。
- Spark on Amazon EMR バージョン 6.6.0、6.7.0、6.8.0 で DynamoDB コネクタを使用すると、テーブルから何を読み込んでも空の結果が返されます。この状況は、入力分割が空でないデータを



参照している場合でも変わりません。Amazon EMR リリース 6.9.0 では、この問題が修正されています。

- Amazon EMR 6.9.0 では、Spark SQL を使用してデータを読み取る際の Apache Hudi による Lake Formation ベースのアクセスコントロールに対する限定的なサポートが追加されています。Spark SQL を使用する SELECT クエリがサポートされます。サポートは、列レベルのアクセス制御に限定されます。詳細については、「[Hudi と Lake Formation](#)」を参照してください。
- Amazon EMR 6.9.0 を使用して、[ノードラベル](#)を有効にした Hadoop クラスターを作成すると、[YARN メトリクス API](#) はデフォルトのパーティションではなく、すべてのパーティションの集計情報を返します。詳細については、「[YARN-11414](#)」を参照してください。
- Amazon EMR リリース 6.9.0 では、Trino を、Java 17 を使用するバージョン 398 にアップデートしました。Amazon EMR 6.8.0 で以前サポートされていた Trino のバージョンは Java 11 で動作する Trino 388 でした。この変更の詳細については、Trino のブログの「[Trino updates to Java 17](#)」を参照してください。
- このリリースでは、Apache BigTop と EC2 クラスターの Amazon EMR の起動シーケンス間のタイミングシーケンスの不一致の問題が修正されています。このタイミングシーケンスの不一致は、システムが 2 つ以上の操作を適切なシーケンスで実行するのではなく、同時に実行しようとした場合に発生します。その結果、特定のクラスター構成でインスタンスの起動がタイムアウトし、クラスターの起動時間が遅くなりました。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

#### Note

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または [ListReleaseLabels](#) API あるいは [list-release-labels](#) CLI アクションを使用するとき Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			連邦)、カナダ (中部)、イスラエル (テルアビブ)
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202307.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 1 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

## リリース 6.8.0

次のリリースノートには、Amazon EMR リリース 6.8.0 に関する情報が含まれています。6.7.0 からの変更が含まれています。

### 新機能

- Amazon EMR ステップ機能が Apache Livy エンドポイントと JDBC/ODBC クライアントをサポートするようになりました。詳細については、「[Configure runtime roles for Amazon EMR steps](#)」を参照してください。
- Amazon EMR リリース 6.8.0 には Apache HBase リリース 2.4.12 が付属しています。この HBase リリースでは、HBase テーブルのアーカイブと削除の両方を行うことができます。Amazon S3 アーカイブプロセスでは、すべてのテーブルファイルの名前がアーカイブディレクトリに変更されます。このプロセスにはコストがかかり、時間を要する場合があります。これで、アーカイブプロセスをスキップして、大きなテーブルをすばやく削除することができます。詳細については、「[HBase シェルを使用する](#)」を参照してください。

### 既知の問題

- Hadoop 3.3.3 では YARN ([YARN-9608](#)) に変更が加えられ、アプリケーションが完了するまでコンテナが実行されていたノードは廃止状態のままになります。この変更により、シャッフルデータなどのローカルデータが失われることはなく、ジョブを再実行する必要もなくなります。Amazon EMR 6.8.0 および 6.9.0 の場合、この方法では、マネージドスケーリングが有効になっているかどうかにかかわらず、クラスターのリソースが十分に活用されない可能性もあります。

[Amazon EMR 6.10.0](#) では、`yarn-site.xml` の `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` の値を `false` に設定することでこの問題を回避できます。Amazon EMR リリース 6.11.0 以降および 6.8.1、6.9.1、6.10.1 では、この問題を解決するために設定がデフォルトで `false` にセットされています。

### 変更、拡張、解決した問題

- Amazon EMR リリース 6.5.0、6.6.0、または 6.7.0 が Apache Spark シェルを介して Apache Phoenix テーブルを読み取ると、Amazon EMR は `NoSuchMethodError` を生成しました。Amazon EMR リリース 6.8.0 では、この問題が修正されています。



- Amazon EMR リリース 6.8.0 には [Apache Hudi 0.11.1](#) が付属していますが、Amazon EMR 6.8.0 クラスターは Hudi 0.12.0 のオープンソース hudi-spark3.3-bundle\_2.12 と互換性があります。
- Amazon EMR リリース 6.8.0 には、Apache Spark 3.3.0 が付属しています。この Spark リリースでは、Apache Log4j 2 と log4j2.properties ファイルを使用して Spark プロセス内の Log4j を設定します。クラスターで Spark を使用するか、カスタム設定パラメータを使用して EMR クラスターを作成し、Amazon EMR リリース 6.8.0 にアップグレードする場合は、Apache Log4j 2 の新しい spark-log4j2 設定分類とキー形式に移行する必要があります。詳細については、「[Apache Log4j 1.x から Log4j 2.x への移行](#)」を参照してください。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

#### Note

このリリースは、さらに 1 つのパッチリリースが続いているため、AMI の自動更新は行われなくなりました。パッチリリースは 2 番目の小数点の後の数字 (6.8.**1**) で示されます。最新のパッチリリースを使用しているかどうかを確認するには、「[リリースガイド](#)」で利用可能なリリースを確認するか、コンソールでクラスターを作成するとき、または [ListReleaseLabels](#) API あるいは [list-release-labels](#) CLI アクションを使用するときに Amazon EMR リリースドロップダウンを確認してください。新しいリリースに関する最新情報を入手するには、「[最新情報](#)」ページの RSS フィードにサブスクライブしてください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイ

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			<p>オ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)</p>

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アジアパシフィック (メルボルン)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			ン)、中東 (アラブ首長国連邦)、カナダ (中部)
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202307.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 119.1	4.14.301	2023 年 2 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022004.0	4.14.294	2022 年 11 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 912.1	4.14.291	2022 年 9 月 6 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

## 既知の問題

- Spark on Amazon EMR バージョン 6.6.0、6.7.0、6.8.0 で DynamoDB コネクタを使用すると、テーブルから何を読み込んでも空の結果が返されます。この状況は、入力分割が空でないデータを参照している場合でも変わりません。これは、Spark 3.2.0 が `spark.hadoopRDD.ignoreEmptySplits` をデフォルトで `true` に設定しているためです。回避策として、明示的に `spark.hadoopRDD.ignoreEmptySplits` を `false` に設定します。Amazon EMR リリース 6.9.0 では、この問題が修正されています。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、Log4j2 プロパティのファイル名パターン設定が正しくないため、SecretAgent および RecordServer サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

リリースタイムラインの詳細については、「[変更ログ](#)」を参照してください。

## リリース 6.7.0

次のリリースノートには、Amazon EMR リリース 6.7.0 に関する情報が含まれています。6.6.0 からの変更が含まれています。

初回リリース日: 2022 年 7 月 15 日

### 新機能

- Amazon EMR は Apache Spark 3.2.1、Apache Hive 3.1.3、HUDI 0.11、PrestoDB 0.272、Trino 0.378 をサポートするようになりました。
- EC2 クラスター上の Amazon EMR 用の EMR ステップ (Spark、Hive) による IAM ロールおよび Lake Formation ベースのアクセスコントロールをサポートします。
- Apache Ranger 対応クラスターでの Apache Spark データ定義ステートメントをサポートします。これには、Apache Ranger 対応クラスターで Apache Hive メタデータを読み書きする Trino アプリケーションのサポートが含まれるようになりました。詳細については、「[Enable federated governance using Trino and Apache Ranger on Amazon EMR](#)」を参照してください。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
			( 米国東部 )、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 207.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023119.1	4.14.301	2023 年 2 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 004.0	4.14.294	2022 年 11 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 912.1	4.14.291	2022 年 10 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 719.0	4.14.287	2022 年 8 月 10 日	us-west-1 , eu-west-3 , eu-north-1 , ap-south-1 , me-south-1

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 606.1	4.14.281	2022 年 7 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

## 既知の問題

- Amazon EMR リリース 6.5.0、6.6.0、または 6.7.0 が Apache Spark シェルを介して Apache Phoenix テーブルを読み取ると、Amazon EMR が誤った `Hbase.compat.version` を使用しているため、`NoSuchMethodError` が発生します。Amazon EMR リリース 6.8.0 では、この問題が修正されています。
- Spark on Amazon EMR バージョン 6.6.0、6.7.0、6.8.0 で DynamoDB コネクタを使用すると、テーブルから何を読み込んでも空の結果が返されます。この状況は、入力分割が空でないデータを参照している場合でも変わりません。これは、Spark 3.2.0 が `spark.hadoopRDD.ignoreEmptySplits` をデフォルトで `true` に設定しているためです。回避策として、明示的に `spark.hadoopRDD.ignoreEmptySplits` を `false` に設定します。Amazon EMR リリース 6.9.0 では、この問題が修正されています。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、`Log4j2` プロパティのファイル名パターン設定が正しくないため、`SecretAgent` および `RecordServer` サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- GetClusterSessionCredentials API は、Amazon EMR 6.7 以前で実行されるクラスターではサポートされていません。

## リリース 6.6.0

次のリリースノートには、Amazon EMR リリース 6.6.0 に関する情報が含まれています。6.5.0 からの変更が含まれています。

初回リリース日: 2022 年 5 月 9 日

ドキュメント更新日: 2022 年 6 月 15 日

### 新機能

- Amazon EMR 6.6 は、Apache Spark 3.2、Apache Spark RAPIDS 22.02、CUDA 11、Apache Hudi 0.10.1、Apache Iceberg 0.13、Trino 0.367、PrestoDB 0.267 をサポートするようになりました。
- Amazon EMR 5.36 以上、6.6 以上、または 7.0 以上の最新のパッチリリースでクラスターを起動すると、Amazon EMR はデフォルトの Amazon EMR AMI に最新の Amazon Linux 2023 または Amazon Linux 2 リリースを使用します。詳細については、「[Amazon EMR にデフォルトの Amazon Linux AMI を使用する](#)」を参照してください。

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 223.0	4.14.336	2024 年 3 月 8 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud

OsReleaseLabel (Amazon Linux バージョン)	Amazon Linux カーネルバージョン	利用可能日	サポートされるリージョン
			( 米国東部 )、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024131.0	4.14.336	2024 年 2 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 124.0	4.14.336	2024 年 2 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2024 109.0	4.14.334	2024 年 1 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 218.0	4.14.330	2024 年 1 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.202306.0	4.14.330	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 116.0	4.14.328	2023 年 12 月 11 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023101.0	4.14.327	2023 年 11 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023020.1	4.14.326	2023 年 11 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023012.1	4.14.326	2023 年 10 月 26 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 926.0	4.14.322	2023 年 10 月 19 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)、AWS GovCloud (米国西部)、AWS GovCloud (米国東部)、中国 (北京)、中国 (寧夏)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 906.0	4.14.322	2023 年 10 月 4 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 822.0	4.14.322	2023 年 8 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.20230808.0	4.14.320	2023 年 8 月 24 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 727.0	4.14.320	2023 年 8 月 14 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (フランクフルト)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 719.0	4.14.320	2023 年 8 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、欧州 (ストックホルム)、欧州 (ミラノ)、欧州 (スペイン)、欧州 (フランクフルト)、欧州 (チューリッヒ)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ハイデラバード)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アジアパシフィック (ジャカルタ)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)、中東 (アラブ首長国連邦)、カナダ (中部)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 628.0	4.14.318	2023 年 7 月 12 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 612.0	4.14.314	2023 年 6 月 23 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 504.1	4.14.313	2023 年 5 月 16 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 418.0	4.14.311	2023 年 5 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 404.1	4.14.311	2023 年 4 月 18 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2023 404.0	4.14.311	2023 年 4 月 10 日	米国東部 (バージニア北部)、欧州 (パリ)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 320.0	4.14.309	2023 年 3 月 30 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 307.0	4.14.305	2023 年 3 月 15 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023 207.0	4.14.304	2023 年 2 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2023119.1	4.14.301	2023 年 2 月 3 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 210.1	4.14.301	2023 年 12 月 22 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)



OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 103.3	4.14.296	2022 年 12 月 5 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 004.0	4.14.294	2022 年 11 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 912.1	4.14.291	2022 年 10 月 7 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)
2.0.2022 805.0	4.14.287	2022 年 8 月 30 日	us-west-1

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 719.0	4.14.287	2022 年 8 月 10 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 426.0	4.14.281	2022 年 6 月 10 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

OsRelease Label (Amazon Linux バージョン)	Amazon Linux カーネル バージョン	利用可能日	サポートされるリージョン
2.0.2022 406.1	4.14.275	2022 年 5 月 2 日	米国東部 (バージニア北部)、米国東部 (オハイオ)、米国西部 (北カリフォルニア)、米国西部 (オレゴン)、カナダ (中部)、欧州 (ストックホルム)、欧州 (アイルランド)、欧州 (ロンドン)、欧州 (パリ)、欧州 (フランクフルト)、欧州 (ミラノ)、アジアパシフィック (香港)、アジアパシフィック (ムンバイ)、アジアパシフィック (ジャカルタ)、アジアパシフィック (東京)、アジアパシフィック (ソウル)、アジアパシフィック (大阪)、アジアパシフィック (シンガポール)、アジアパシフィック (シドニー)、アフリカ (ケープタウン)、南米 (サンパウロ)、中東 (バーレーン)

- Amazon EMR 6.6 以降では、Log4j 1.x と Log4j 2.x を使用するアプリケーションは、それぞれ Log4j 1.2.17 (またはそれ以上) と Log4j 2.17.1 (またはそれ以上) を使用するようにアップグレードされます。CVE の問題を軽減するために提供されている [ブートストラップアクション](#) を使用する必要はなくなります。

- [マネージドスケールリング] Spark シャッフルデータマネージドスケールリング最適化 - Amazon EMR バージョン 5.34.0 以降、および EMR バージョン 6.4.0 以降では、マネージドスケールリングは Spark シャッフルデータ (Spark が特定の操作を実行するためにパーティション間で再分配するデータ) 対応になりました。シャッフル操作の詳細については、「Amazon EMR 管理ガイド」および「[Spark のプログラミングガイド](#)」の「[Amazon EMR での EMR マネージドスケールリングの使用](#)」を参照してください。
- Amazon EMR 5.32.0 および 6.5.0 以降、Apache Spark の動的エグゼキューターサイズ設定はデフォルトで有効になっています。この機能をオンまたはオフにするには、`spark.yarn.heterogeneousExecutors.enabled` 設定パラメータを使用できます。

## 変更、拡張、解決した問題

- Amazon EMR は、EMR のデフォルト AMI オプションを使用し、Apache Hadoop、Apache Spark、Apache Hive などの一般的なアプリケーションのみをインストールするクラスターの場合、クラスターの起動時間を平均で最大 80 秒短縮します。

## 既知の問題

- Amazon EMR リリース 6.5.0、6.6.0、または 6.7.0 が Apache Spark シエルを介して Apache Phoenix テーブルを読み取ると、Amazon EMR が誤った `Hbase.compat.version` を使用しているため、`NoSuchMethodError` が発生します。Amazon EMR リリース 6.8.0 では、この問題が修正されています。
- Spark on Amazon EMR バージョン 6.6.0、6.7.0、6.8.0 で DynamoDB コネクタを使用すると、テーブルから何を読み込んでも空の結果が返されます。この状況は、入力分割が空でないデータを参照している場合でも変わりません。これは、Spark 3.2.0 が `spark.hadoopRDD.ignoreEmptySplits` をデフォルトで `true` に設定しているためです。回避策として、明示的に `spark.hadoopRDD.ignoreEmptySplits` を `false` に設定します。Amazon EMR リリース 6.9.0 では、この問題が修正されています。
- Trino の長時間稼働クラスターでは、Amazon EMR 6.6.0 は Trino `jvm.config` のガベージコレクションログ記録パラメータを有効にして、ガベージコレクションログからより深い知見を得ることができます。この変更により、`launcher.log (/var/log/trino/launcher.log)` ファイルに多くのガベージコレクションログが追加されます。Amazon EMR 6.6.0 で Trino クラスターを実行している場合、追加されたログが原因で、クラスターを数日間実行した後にノードのディスク容量が不足することがあります。

この問題の回避策は、Amazon EMR 6.6.0 のクラスターを作成またはクローン作成する際に、以下のスクリプトをブートストラップアクションとして実行し、`jvm.config` のガベージコレクションログ記録パラメータを無効にすることです。

```
#!/bin/bash
set -ex
PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の

`spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

- Amazon EMR リリース 5.36.0、および 6.6.0 から 6.9.0 では、`Log4j2` プロパティのファイル名パターン設定が正しくないため、`SecretAgent` および `RecordServer` サービスコンポーネントでログデータが失われる可能性があります。設定が正しくないと、コンポーネントは 1 日に 1 つのログファイルしか生成しません。ローテーション戦略が実行されると、期待どおりに新しいログファイルが生成されず、既存のファイルが上書きされます。回避策として、ブートストラップアクションを使用して 1 時間ごとにログファイルを生成し、ファイル名に自動増分整数を追加してローテーションを処理します。

Amazon EMR 6.6.0 から 6.9.0 のリリースでは、クラスターを起動するときに次のブートストラップアクションを使用します。



```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0 では、クラスターを起動するときに次のブートストラップアクションを使用します。

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

## リリース 5.35.0

これは Amazon EMR リリース 5.35.0 のリリースノートです。

次のリリースノートには、Amazon EMR リリース 5.35.0 に関する情報が含まれています。5.34.0 からの変更が含まれています。

初回リリース日: 2022 年 3 月 30 日

### 新機能

- Log4j 1.x と Log4j 2.x を使用する Amazon EMR リリース 5.35 アプリケーションは、それぞれ Log4j 1.2.17 (またはそれ以上) と Log4j 2.17.1 (またはそれ以上) を使用するようにアップグレードされます。以前のリリースの CVE の問題を軽減するためにブートストラップアクションを使用する必要はなくなります。[CVE-2021-44228 を軽減するためのアプローチ](#) を参照してください。

### 変更、拡張、解決した問題

#### Flink の変更

タイプの変更	説明
アップグレード	<ul style="list-style-type: none"><li>Flink のバージョンを 1.14.2 にアップデートします。</li><li>log4j は 2.17.1 にアップグレードされました。</li></ul>

## Hadoop の変更

タイプの変更	説明
EMR 5.34.0 以降の Hadoop オープンソース バックポート	<ul style="list-style-type: none"> <li>• <a href="#">YARN-10438</a>: ClientRMService #getContainerReport() で null containerId を処理する</li> <li>• <a href="#">YARN-7266</a>: タイムラインサーバーのイベントハンドラスレッドがロックされている</li> <li>• <a href="#">YARN-10438</a>: ファイルが破損しているか、見つからない場合、ATS 1.5 RollingLevelDb は起動に失敗する</li> <li>• <a href="#">HADOOP-13500</a>: 設定プロパティオブジェクトのイテレーションの同期</li> <li>• <a href="#">YARN-10651</a>: . AbstractYarnScheduler.updateNodeResource() の NPE で CapacityScheduler クラッシュ</li> <li>• <a href="#">HDFS-12221</a>: の xerces を置き換える XmlEditsVisitor</li> <li>• <a href="#">HDFS-16410</a>: での安全でない Xml 解析 OfflineEditsXmlLoader</li> </ul>
Hadoop の変更と修正	<ul style="list-style-type: none"> <li>• KMS と HttpFS で使用されている Tomcat が 8.5.75 にアップグレードされました。</li> <li>• FileSystemOptimizedCommitterV2 では、コミッターの作成中に定義された commitJob 出力パスに成功マーカが書き込まれました。commitJob とタスクレベルの出力パスは異なる場合があるため、マニフェストファイルで定義されているものを使用するようにパスが修正されました。これにより、Hive ジョブで、動的パーティションや UNION ALL などの操作を実行すると、成功マーカが正しく書き込まれるようになりました。</li> </ul>

## Hive の変更点

タイプの変更	説明
<p>Hive がオープンソース <a href="#">リリース 2.3.9</a> にアップグレードされました (これらの JIRA の修正を含む)</p>	<ul style="list-style-type: none"> <li>• HiveConf.java の <a href="#">HIVE-17155</a> : findConfFile() には conf パスにいくつかの問題があります</li> <li>• <a href="#">HIVE-24797</a>: Avro スキーマを解析するときのデフォルト値の検証を無効にする</li> <li>• <a href="#">HIVE-21563</a>: 1 registerAllFunctions回無効にして Table#getEmptyTable performance を改善</li> <li>• <a href="#">HIVE-18147</a>: java.net でテストが失敗する可能性があります。BindException: アドレスは既に使用されています</li> <li>• <a href="#">HIVE-24608</a>: Hive 2.3.x の HMS クライアントで get_table に切り替える</li> <li>• <a href="#">HIVE-21200</a>: ベクトル化 - Parquet UnsupportedOperationException 用の java.lang. をスローする日付列</li> <li>• <a href="#">HIVE-19228</a>: commons-httpclient 3.x の使用を削除する</li> </ul>
<p>EMR 5.34.0 以降の Hive オープンソースバックポート</p>	<ul style="list-style-type: none"> <li>• <a href="#">HIVE-19990</a>: 結合条件に間隔のリテラルを含むクエリが失敗する</li> <li>• <a href="#">HIVE-25824</a>: branch-2.3 を log4j 2.17.0 にアップグレードする</li> <li>• <a href="#">TEZ-4062</a>: 投機的試行のスケジューリングはタスクが完了したら中止する必要がある</li> <li>• <a href="#">TEZ-4108</a>: 投機的実行の競合状態中の NullPointerException</li> <li>• <a href="#">TEZ-3918</a>: tez.task.log.level を設定しても機能しない</li> </ul>
<p>Hive のアップグレードと修正</p>	<ul style="list-style-type: none"> <li>• Log4j バージョンの 2.17.1 へのアップグレード</li> </ul>

タイプの変更	説明
	<ul style="list-style-type: none"> <li>• ORC バージョンの 1.4.3 へのアップグレード</li> <li>• のペナルティスレッドによるFixED ShuffleScheduler</li> </ul>
新機能	<ul style="list-style-type: none"> <li>• Hive クエリを AM ログに出力する機能が追加されました。これはデフォルトでは無効になっています。フラグ/設定: <code>tez.am.emr.print.hive.query.in.log</code> 。ステータス (デフォルト): FALSE。</li> </ul>

## Oozie の変更

タイプの変更	説明
EMR 5.34.0 以降の Oozie オープンソースバックポート	<ul style="list-style-type: none"> <li>• <a href="#">OOZIE-3652</a>: Oozie ランチャーは <code>NoSuchFileException</code> 、発生時にディレクトリのリストを再試行する必要があります</li> </ul>

## Pig の変更

タイプの変更	説明
アップグレード	<ul style="list-style-type: none"> <li>• <code>log4j</code> は 1.2.17 にアップグレードされました。</li> </ul>

## 既知の問題

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。

- 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
- 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## リリース 5.34.0

次のリリースノートには、Amazon EMR リリース 5.34.0 に関する情報が含まれています。5.33.1 からの変更が含まれています。

初回リリース日: 2022 年 1 月 20 日

更新リリース日: 2022 年 3 月 21 日

### 新機能

- [マネージドスケールリング] Spark シャッフルデータマネージドスケールリング最適化 - Amazon EMR バージョン 5.34.0 以降、および EMR バージョン 6.4.0 以降では、マネージドスケールリングは Spark シャッフルデータ (Spark が特定の操作を実行するためにパーティション間で再分配するデータ) 対応になりました。シャッフル操作の詳細については、「Amazon EMR 管理ガイド」および「[Spark のプログラミングガイド](#)」の「[Amazon EMR での EMR マネージドスケールリングの使用](#)」を参照してください。
- [Hudi] Hudi の設定を簡素化するための改良点。オプティミスティック同時実行制御をデフォルトで無効にしました。

### 変更、拡張、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

- 以前は、マルチマスタークラスターでリソースマネージャーを手動で再起動すると、Zookeeper などの Amazon EMR のクラスター上のデーモンが、Zookeeper znode ファイル内の以前に廃止された、または失われたすべてのノードを再ロードしていました。これにより、特定の状況でデフォルトの制限を超えることがありました。Amazon EMR では、1 時間以上経過した、廃止された、または失われたノードレコードが Zookeeper ファイルから削除されるようになり、内部制限も引き上げられました。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Zeppelin をバージョン 0.10.0 にアップグレードしました。
- Livy Fix - 0.7.1 にアップグレード
- Spark のパフォーマンスの向上 - EMR 5.34.0 で特定の Spark 設定値がオーバーライドされると、異種工グゼキューターが無効になります。
- WebHDFS と HttpFS サーバーはデフォルトで無効になっています。Hadoop 設定 `dfs.webhdfs.enabled` を使用して WebHDFS を再度有効にすることができます。HttpFS サーバーを起動するには、`sudo systemctl start hadoop-httpfs` を使用します。

## 既知の問題

- Livy ユーザー偽装で使用される Amazon EMR Notebooks 機能は、HttpFS がデフォルトで無効になっているため機能しません。この場合、EMR notebooks は Livy 偽装が有効になっているクラスターに接続できません。回避策は、`sudo systemctl start hadoop-https` を使用して EMR notebooks をクラスターに接続する前に、HttpFS サーバーを起動することです。
- Apache Hadoop HttpFS サーバーはデフォルトで無効になるため、Hue クエリは Amazon EMR 6.4.0 では機能しません。Amazon EMR 6.4.0 で Hue を使用するには、`sudo systemctl start hadoop-https` を使用して Amazon EMR プライマリノードで HttpFS サーバーを手動で起動するか、または [Amazon EMR のステップを使用](#) します。
- Livy ユーザー偽装で使用される Amazon EMR Notebooks 機能は、HttpFS がデフォルトで無効になっているため機能しません。この場合、EMR notebooks は Livy 偽装が有効になっているクラスターに接続できません。回避策は、`sudo systemctl start hadoop-https` を使用して EMR notebooks をクラスターに接続する前に、HttpFS サーバーを起動することです。
- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## リリース 6.5.0

次のリリースノートには、Amazon EMR リリース 6.5.0 に関する情報が含まれています。6.4.0 からの変更が含まれています。

初回リリース日: 2022 年 1 月 20 日

更新リリース日: 2022 年 3 月 21 日

## 新機能

- [マネージドスケールリング] Spark シャッフルデータマネージドスケールリング最適化 - Amazon EMR バージョン 5.34.0 以降、および EMR バージョン 6.4.0 以降では、マネージドスケールリングは Spark シャッフルデータ (Spark が特定の操作を実行するためにパーティション間で再分配するデータ) 対応になりました。シャッフル操作の詳細については、「Amazon EMR 管理ガイド」および「[Spark のプログラミングガイド](#)」の「[Amazon EMR での EMR マネージドスケールリングの使用](#)」を参照してください。
- Amazon EMR 5.32.0 および 6.5.0 以降、Apache Spark の動的エグゼキューターサイズ設定はデフォルトで有効になっています。この機能をオンまたはオフにするには、`spark.yarn.heterogeneousExecutors.enabled` 設定パラメータを使用できます。
- 巨大な分析データセット用の Apache Iceberg オープンテーブル形式のサポート。
- ranger-trino-plugin 2.0.1-amzn-1 のサポート
- toree 0.5.0 のサポート

## 変更、拡張、解決した問題

- Amazon EMR 6.5 リリースバージョンは Apache Iceberg 0.12.0 をサポートするようになり、Apache Spark 用 Amazon EMR ランタイム、Presto 用 Amazon EMR ランタイム、Apache Hive 用 Amazon EMR ランタイムによりランタイムが改善されています。
- [Apache Iceberg](#) は、Amazon S3 の大規模データセット用のオープンテーブル形式であり、大きなテーブルに対する高速のクエリパフォーマンス、アトミックコミット、同時書き込み、および SQL 互換テーブルの進化を提供します。EMR 6.5 では、Apache Spark 3.1.2 を Iceberg テーブル形式で使用できます。
- Apache Hudi 0.9 では Spark SQL DDL と DML のサポートが追加されています。これにより、SQL ステートメントだけを使用して Hudi テーブルを作成し、アップサートできるようになりました。Apache Hudi 0.9 では、クエリ側とライター側のパフォーマンスも向上しています。
- Apache Hive 用 Amazon EMR ランタイムは、ステージング操作中の名前変更操作を削除することで Amazon S3 での Apache Hive のパフォーマンスを向上させ、テーブルの修復に使用されるメタストアチェック (MSCK) コマンドのパフォーマンスを向上させます。



## 既知の問題

- Amazon EMR リリース 6.5.0、6.6.0、または 6.7.0 が Apache Spark シェルを介して Apache Phoenix テーブルを読み取ると、Amazon EMR が誤った `Hbase.compat.version` を使用しているため、`NoSuchMethodError` が発生します。Amazon EMR リリース 6.8.0 では、この問題が修正されています。
- 高可用性 (HA) の Hbase バンドルクラスターは、デフォルトのボリュームサイズとインスタンスタイプではプロビジョニングに失敗します。この問題の回避策は、ルートボリュームのサイズを増やすことです。
- Apache Oozie で Spark アクションを使用するには、以下の設定を Oozie workflow.xml ファイルに追加する必要があります。そうしないと、Hadoop や EMRFS などのいくつかの重要なライブラリが、Oozie が起動する Spark エグゼキューターのクラスパスから失われてしまいます。

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## リリース 6.4.0

次のリリースノートには、Amazon EMR リリース 6.4.0 に関する情報が含まれています。6.3.0 からの変更が含まれています。

初回リリース日: 2021 年 9 月 20 日

更新リリース日: 2022 年 3 月 21 日

### サポートされているアプリケーション

- AWS SDK for Java バージョン 1.12.31
- CloudWatch シンクバージョン 2.2.0
- DynamoDB Connector バージョン 4.16.0
- EMRFS バージョン 2.47.0
- Amazon EMR Goodies バージョン 3.2.0
- Amazon EMR Kinesis Connector バージョン 3.5.0
- Amazon EMR Record Server バージョン 2.1.0
- Amazon EMR Scripts バージョン 2.5.0
- Flink バージョン 1.13.1
- Ganglia バージョン 3.7.2
- AWS Glue Hive メタストアクライアントバージョン 3.3.0
- Hadoop バージョン 3.2.1-amzn-4
- HBase バージョン 2.4.4-amzn-0
- HBase-operator-tools 1.1.0
- HCatalog バージョン 3.1.2-amzn-5
- Hive バージョン 3.1.2-amzn-5
- Hudi バージョン 0.8.0-amzn-0
- Hue バージョン 4.9.0
- Java JDK バージョン Corretto-8.302.08.1 (ビルド 1.8.0\_302-b08)
- JupyterHub バージョン 1.4.1
- Livy バージョン 0.7.1-incubating

- MXNet バージョン 1.8.0
- Oozie バージョン 5.2.1
- Phoenix バージョン 5.1.2
- Pig バージョン 0.17.0
- Presto バージョン 0.254.1-amzn-0
- Trino バージョン 359
- Apache Ranger KMS (マルチマスター透過的暗号化) バージョン 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker Spark SDK バージョン 1.4.1
- Scala バージョン 2.12.10 (OpenJDK 64 ビットサーバー VM、Java 1.8.0\_282)
- Spark バージョン 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop バージョン 1.4.7
- TensorFlow バージョン 2.4.1
- tez バージョン 0.9.2
- Zeppelin バージョン 0.9.0
- Zookeeper バージョン 3.5.7
- コネクタおよびドライバー: DynamoDB Connector 4.16.0

## 新機能

- [マネージドスケーリング] Spark シャッフルデータマネージドスケーリング最適化 - Amazon EMR バージョン 5.34.0 以降、および EMR バージョン 6.4.0 以降では、マネージドスケーリングは Spark シャッフルデータ (Spark が特定の操作を実行するためにパーティション間で再分配するデータ) 対応になりました。シャッフル操作の詳細については、「[Amazon EMR 管理ガイド](#)」および「[Spark のプログラミングガイド](#)」の「[Amazon EMR での EMR マネージドスケーリングの使用](#)」を参照してください。
- Apache Ranger 対応の Amazon EMR クラスターでは、Apache Spark SQL を使用して、INSERT INTO、INSERT OVERWRITE、および ALTER TABLE で Apache Hive メタストアテーブルでデータの挿入または更新を実行できます。Spark SQL で ALTER TABLE を使用する場合、パーティションの場所はテーブルの場所の子ディレクトリである必要があります。Amazon EMR は現在、

パーティションの場所がテーブルの場所と異なるパーティションへのデータの挿入をサポートしていません。

- PrestoSQL は [Trino に名称変更](#)されています。
- Hive: LIMIT 句を使用した簡単な SELECT クエリの実行は、LIMIT 句に記述されているレコード数がフェッチされたらすぐにクエリの実行を停止すると、高速化されます。簡単な SELECT クエリは、GROUP BY/ORDER BY 句のないクエリ、またはリデューサーステージのないクエリです。例えば SELECT \* from <TABLE> WHERE <Condition> LIMIT <Number> です。

## Hudi の同時実行制御

- Hudi はオプティミスティック同時実行制御 (OCC) をサポートするようになりました。この OCC を UPSERT や INSERT などの書き込みオペレーションで利用して、複数のライターから同じ Hudi テーブルへの変更を許可できます。これはファイルレベルの OCC であるため、2 つのコミット (またはライター) は、その変更が競合しなければ、同じテーブルに書き込むことができます。詳細については、「[Hudi Concurrency Control](#)」を参照してください。
- Amazon EMR クラスターには Zookeeper がインストールされており、OCC のロックプロバイダーとして利用できます。この機能を使いやすくするために、Amazon EMR クラスターには次のプロパティが事前設定されています。

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

OCC を有効にするには、Hudi ジョブオプションを使用して、または Amazon EMR 設定 API を使用してクラスターレベルで、次のプロパティを設定する必要があります。

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table Name is a good option)
```

## Hudi モニタリング: Hudi メトリクスを報告するための Amazon CloudWatch 統合

- Amazon EMR は、Amazon への Hudi メトリクスの発行をサポートしています CloudWatch。これを有効にするには、次の必要な設定を行います。

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- 変更できるオプションの Hudi 設定を以下に示します。

設定	説明	値
hoodie.metrics.cloudwatch.report.period.seconds	Amazon にメトリクスを報告する頻度 (秒単位) CloudWatch	デフォルト値は 60 秒で、Amazon が提供するデフォルトの 1 分間の解像度でも問題ありません。CloudWatch
hoodie.metrics.cloudwatch.metric.prefix	各メトリクス名に追加するプレフィックス	デフォルト値は空です (プレフィックスなし)
hoodie.metrics.cloudwatch.namespace	メトリクスが公開される Amazon CloudWatch 名前空間	デフォルト値は Hudi です
hoodie.metrics.cloudwatch.maxDatumsPerRequest	Amazon への 1 つのリクエストに含めるデータムの最大数 CloudWatch	デフォルト値は 20 で、Amazon の CloudWatch デフォルトと同じです。

## Amazon EMR Hudi の設定のサポートと改善

- EMR 設定 API と再設定機能を利用して、クラスターレベルで Hudi 設定を構成できるようになりました。Spark、Hive などの他のアプリケーションに似た `/etc/hudi/conf/hudi-defaults.conf` を使用して、新しいファイルベースの設定のサポートが導入されました。EMR では、ユーザーエクスペリエンスを向上させるためにデフォルト設定はほとんどありません。

— `hoodie.datasource.hive_sync.jdbcurl` はクラスター Hive サーバ URL に設定されるため、指定する必要がなくなりました。これは、Spark クラスターモードでジョブを実行する場合に特に便利です。この場合、以前は Amazon EMR マスター IP を指定する必要がありました。

— HBase 固有の設定。Hbase のインデックスを Hudi で使用する場合に役立ちます。

- 同時実行制御で説明されているように、Zookeeper はプロバイダー固有の設定をロックします。これにより、オプティミスティック同時実行制御 (OCC) の使用が容易になります。
- 渡す必要がある設定の数を減らし、可能な場合は自動的に推測するために、追加の変更が導入されました。
  - `partitionBy` キーワードは、パーティション列を指定するために使用できます。
  - Hive Sync を有効にすると、`HIVE_TABLE_OPT_KEY`, `HIVE_PARTITION_FIELDS_OPT_KEY`, `HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY` を渡すのは必須ではなくなります。これらの値は、Hudi テーブル名とパーティションフィールドから推測できます。
  - `KEYGENERATOR_CLASS_OPT_KEY` を渡すのは必須ではなく、より単純な `SimpleKeyGenerator` および `ComplexKeyGenerator` のケースから推測できます。

## Hudi Caveats

- Hudi では、読み取り時マージ (MoR) テーブルおよびブートストラップテーブルの Hive でのベクトル化された実行をサポートしていません。例えば、`hive.vectorized.execution.enabled` が `true` に設定されている場合、Hudi リアルタイムテーブルで `count(*)` は失敗します。回避策として、`hive.vectorized.execution.enabled` を `false` に設定して、ベクトル化された読み取りを無効にすることができます。
- マルチライターサポートは、Hudi ブートストラップ機能とは互換性がありません。
- Flink Streamer と Flink SQL は、このリリースでは実験的な機能です。これらの機能は、実稼働環境での使用はお勧めしません。

## 変更点、機能強化、解決した問題

これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

- 以前は、マルチマスタークラスターでリソースマネージャーを手動で再起動すると、Zookeeper などの Amazon EMR のクラスター上のデーモンが、Zookeeper znode ファイル内の以前に廃止された、または失われたすべてのノードを再ロードしていました。これにより、特定の状況でデフォルトの制限を超えることがありました。Amazon EMR では、1 時間以上経過した、廃止された、または失われたノードレコードが Zookeeper ファイルから削除されるようになり、内部制限も引き上げられました。

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Apache YARN タイムラインサーバーのバージョン 1 および 1.5 のパフォーマンスの問題を修正するためのクラスターの設定

Apache YARN タイムラインサーバーのバージョン 1 および 1.5 では、特に Amazon EMR のデフォルト設定である `yarn.resourcemanager.system-metrics-publisher.enabled=true` を使用する場合に、非常にアクティブで大規模な EMR クラスターでパフォーマンスの問題が発生する可能性があります。オープンソースの YARN タイムラインサーバー v2 では、YARN タイムラインサーバーのスケラビリティに関連するパフォーマンスの問題が解決されています。

この問題の他の回避策には、次のものがあります。

- `yarn-site.xml` での `yarn.resourcemanager.system-metrics-publisher.enabled=false` の設定。
- クラスターの作成時にこの問題の修正を有効にします (以下を参照)。

次の Amazon EMR リリースには、この YARN タイムラインサーバーのパフォーマンス問題の修正が含まれています。

EMR 5.30.2、5.31.1、5.32.1、5.33.1、5.34.x、6.0.1、6.1.1、6.2.1、6.3.1、6.4.x



上記に示されている Amazon EMR リリースで修正を有効にするには、[aws emr create-cluster コマンドのパラメータ](#) `--configurations file://./configurations.json` で渡される設定の JSON ファイルで、以下のプロパティを `true` に設定します。または、[再構成コンソール UI](#) を使用して修正を有効にします。

configurations.json ファイルの内容の例

```
[
{
  "Classification": "yarn-site",
  "Properties": {
    "yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
      "true",
    "yarn.resourcemanager.system-metrics-publisher.enabled": "true"
  },
  "Configurations": []
}
]
```

- WebHDFS と HttpFS サーバーはデフォルトで無効になっています。Hadoop 設定 `dfs.webhdfs.enabled` を使用して WebHDFS を再度有効にすることができます。HttpFS サーバーを起動するには、`sudo systemctl start hadoop-httpfs` を使用します。
- Amazon Linux リポジトリでは、現在、HTTPS がデフォルトで有効になります。Amazon S3 VPC ポリシーを使用して特定のバケットへのアクセスを制限する場合は、新しい Amazon Linux バケット ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` をポリシーに追加する (`$region` を、エンドポイントがあるリージョンに置き換える) 必要があります。詳細については、AWS ディスカッションフォーラムのこのトピックを参照してください。[Announcement: Amazon Linux 2 now supports the ability to use HTTPS while connecting to package repositories](#)
- Hive: 最後のジョブに HDFS でスクラッチディレクトリを使用できるようにすると、書き込みクエリのパフォーマンスが向上します。最後のジョブの一時データは Amazon S3 ではなく HDFS に書き込まれ、データは Amazon S3 デバイス間ではなく HDFS から最後のテーブルの場所 (Amazon S3) に移動されるため、パフォーマンスが向上します。
- Hive: Glue メタストアパーティショニングにより、クエリのコンパイル時間が最大 2.5 倍に改善されています。
- デフォルトでは、組み込みの UDF が Hive から Hive メタストアサーバーに渡されると、Glue は限定された式演算子しかサポートしていないため、組み込みの UDF のサブセットのみが Glue メタストアに渡されます。`hive.glue.partition.pruning.client=true`



を設定した場合、パーティションのプルーニングはすべてクライアント側で行われます。hive.glue.partition.pruning.server=true を設定した場合、パーティションのプルーニングはすべてサーバー側で行われます。

## 既知の問題

- Apache Hadoop HttpFS サーバーはデフォルトで無効になるため、Hue クエリは Amazon EMR 6.4.0 では機能しません。Amazon EMR 6.4.0 で Hue を使用するには、`sudo systemctl start hadoop-httpfs` を使用して Amazon EMR プライマリノードで HttpFS サーバーを手動で起動するか、または [Amazon EMR のステップを使用](#) します。
- Livy ユーザー偽装で使用される Amazon EMR Notebooks 機能は、HttpFS がデフォルトで無効になっているため機能しません。この場合、EMR notebooks は Livy 偽装が有効になっているクラスターに接続できません。回避策は、`sudo systemctl start hadoop-httpfs` を使用して EMR notebooks をクラスターに接続する前に、HttpFS サーバーを起動することです。
- Amazon EMR バージョン 6.4.0 では、Phoenix は Phoenix コネクタコンポーネントをサポートしていません。
- Apache Oozie で Spark アクションを使用するには、以下の設定を Oozie workflow.xml ファイルに追加する必要があります。そうしないと、Hadoop や EMRFS などのいくつかの重要なライブラリが、Oozie が起動する Spark エグゼキューターのクラスパスから失われてしまいます。

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## リリース 5.32.0

次のリリースノートには、Amazon EMR リリース 5.32.0 に関する情報が含まれています。5.31.0 からの変更が含まれています。

初回リリース日: 2021 年 1 月 8 日

### アップグレード

- Amazon Glue コネクタをバージョン 1.14.0 にアップグレードしました
- Amazon SageMaker Spark SDK をバージョン 1.4.1 にアップグレードしました
- バージョン 1.11.890 AWS SDK for Java にアップグレード
- EMR DynamoDB Connector バージョン 4.16.0 にアップグレードしました
- EMRFS をバージョン 2.45.0 にアップグレードしました
- EMR Log Analytics Metrics をバージョン 1.18.0 にアップグレードしました
- EMR MetricsAndEventsApiGateway クライアントをバージョン 1.5.0 にアップグレードしました
- EMR Record Server をバージョン 1.8.0 にアップグレードしました
- EMR S3 Dist CP をバージョン 2.17.0 にアップグレードしました
- EMR Secret Agent をバージョン 1.7.0 にアップグレードしました
- Flink をバージョン 1.11.2 にアップグレードしました
- Hadoop をバージョン 2.10.1-amzn-0 にアップグレードしました
- Hive をバージョン 2.3.7-amzn-3 にアップグレードしました
- Hue をバージョン 4.8.0 にアップグレードしました
- MXNet をバージョン 1.7.0 にアップグレードしました
- OpenCV をバージョン 4.4.0 にアップグレードしました
- Presto をバージョン 0.240.1-amzn-0 にアップグレードしました
- Spark をバージョン 2.4.7-amzn-0 にアップグレードしました
- バージョン 2.3.1 TensorFlow にアップグレード

## 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- コンポーネントのバージョンをアップグレードしました。
- コンポーネントのバージョンのリストについては、このガイドの「[Amazon EMR リリースについて](#)」を参照してください。

## 新機能

- Amazon EMR 5.32.0 および 6.5.0 以降、Apache Spark の動的エグゼキューターサイズ設定はデフォルトで有効になっています。この機能をオンまたはオフにするには、`spark.yarn.heterogeneousExecutors.enabled` 設定パラメータを使用できます。

- インスタンスメタデータサービス (IMDS) V2 サポートステータス: Amazon EMR 5.23.1、5.27.1、および 5.32 以降のコンポーネントは、すべての IMDS 呼び出しに IMDSv2 を使用します。アプリケーションコードでの IMDS 呼び出しでは、IMDSv1 と IMDSv2 の両方を使用するか、または、セキュリティを強化するために IMDSv2 のみを使用するように IMDS を設定できます。その他の 5.x EMR リリースでは、IMDSv1 を無効にすると、クラスターの起動が失敗します。
- Amazon EMR 5.32.0 以降では、Apache Ranger とネイティブに統合されるクラスターを起動できます。Apache Ranger は、Hadoop プラットフォーム全体で包括的なデータセキュリティを有効化、監視、管理するためのオープンソースフレームワークです。詳細については、「[Apache Ranger](#)」を参照してください。ネイティブ統合により、独自の Apache Ranger を導入して、Amazon EMR できめ細かなデータアクセス制御を実施できます。「Amazon EMR 管理ガイド」の「[Amazon EMR と Apache Ranger を統合する](#)」を参照してください。
- Amazon EMR リリース 5.32.0 は Amazon EMR on EKS をサポートしています。EMR on EKS の開始方法の詳細については、「[Amazon EMR on EKS とは](#)」を参照してください。
- Amazon EMR リリース 5.32.0 は Amazon EMR Studio (プレビュー) をサポートしています。EMR Studio の開始方法の詳細については、「[Amazon EMR Studio \(Preview\)](#)」を参照してください。
- スコープ管理ポリシー: AWS ベストプラクティスに合わせて、Amazon EMR では v2 EMR スコープのデフォルト管理ポリシーを廃止されるポリシーの置き換えとして導入しました。「[Amazon EMR 管理ポリシー](#)」を参照してください。

## 既知の問題

- Amazon EMR 6.3.0 および 6.2.0 のプライベートサブネットクラスターで、Ganglia ウェブ UI にアクセスできません。「アクセス拒否 (403)」というエラーが表示されます。Spark、Hue、Zeppelin、Livy JupyterHub、Tez などの他のウェブ UIs は正常に動作しています。パブリックサブネットクラスターでの Ganglia ウェブ UI アクセスも正常に動作します。この問題を解決するには、`sudo systemctl restart httpd` を使用してプライマリノードで httpd サービスを再起動します。この問題は、Amazon EMR 6.4.0 で修正されています。
- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大

オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
```

```
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

### ⚠ Important

Amazon Linux または Amazon Linux 2 Amazon マシンイメージ (AMI) を実行する EMR クラスターは、デフォルトの Amazon Linux 動作を使用します。再起動が必要な重要かつクリティカルなカーネル更新が自動的にダウンロードされてインストールされることはありません。これは、デフォルトの Amazon Linux AMI を実行している他の Amazon EC2 インスタンスと同じ動作です。Amazon EMR リリースが利用可能になった後に、再起動が必要な新しい Amazon Linux ソフトウェアアップデート (カーネル、NVIDIA、CUDA のアップデートなど) が使用可能になった場合、デフォルトの AMI を実行する EMR クラスターインスタンスで、それらの更新が自動的にダウンロードされてインストールされることはありません。カーネルの更新を取得するには、[Amazon EMR AMI をカスタマイズ](#)して、[最新の Amazon Linux AMI を使用](#)できるようにします。

- AWS Ranger 統合オプションを指定するセキュリティ設定を作成するためのコンソールサポートは、現在 GovCloud リージョンではサポートされていません。セキュリティ設定は CLI を使用して実行できます。「Amazon EMR 管理ガイド」の「[EMR セキュリティ設定を作成する](#)」を参照してください。
- Amazon EMR 5.31.0 AtRestEncryption または 5.32.0 を使用するクラスターで または HDFS 暗号化が有効になっている場合、Hive クエリでは次のランタイム例外が発生します。

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
java.lang.RuntimeException: Hive Runtime Error while closing
operators: java.io.IOException: java.util.ServiceConfigurationError:
org.apache.hadoop.security.token.TokenIdentifier: Provider
org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。



- 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
- 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## リリース 6.2.0

次のリリースノートには、Amazon EMR リリース 6.2.0 に関する情報が含まれています。6.1.0 からの変更が含まれています。

初回リリース日: 2020 年 12 月 9 日

最終更新日: 2021 年 10 月 4 日

サポートされているアプリケーション

- AWS SDK for Java バージョン 1.11.828
- `emr-record-server` バージョン 1.7.0
- Flink バージョン 1.11.2
- Ganglia バージョン 3.7.2
- Hadoop バージョン 3.2.1-amzn-1
- HBase バージョン 2.2.6-amzn-0
- HBase-operator-tools 1.0.0
- HCatalog バージョン 3.1.2-amzn-0
- Hive バージョン 3.1.2-amzn-3
- Hudi バージョン 0.6.0-amzn-1
- Hue バージョン 4.8.0

- JupyterHub バージョン 1.1.0
- Livy バージョン 0.7.0
- MXNet バージョン 1.7.0
- Oozie バージョン 5.2.0
- Phoenix バージョン 5.0.0
- Pig バージョン 0.17.0
- Presto バージョン 0.238.3-amzn-1
- PrestoSQL バージョン 343
- Spark バージョン 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow バージョン 2.3.1
- Zeppelin バージョン 0.9.0-preview1
- Zookeeper バージョン 3.4.14
- コネクタおよびドライバー: DynamoDB Connector 4.16.0

## 新機能

- HBase: コミットフェーズでの名前変更を削除し、永続的 HFile トラッキングを追加しました。「Amazon EMR リリース ガイド」の「[永続的 HFile トラッキング](#)」を参照してください。
- HBase: 「[Create a config that forces to cache blocks on compaction](#)」をバックポートしました。
- PrestoDB: ダイナミックパーティションプルーニングの改善。ルールベースの結合順序は、パーティション分割されていないデータに対して機能します。
- スコープ管理ポリシー: AWS ベストプラクティスに合わせて、Amazon EMR では v2 EMR スコープのデフォルト管理ポリシーを廃止されるポリシーの置き換えとして導入しました。「[Amazon EMR 管理ポリシー](#)」を参照してください。
- インスタンスメタデータサービス (IMDS) V2 サポートステータス: Amazon EMR 6.2 以降では、Amazon EMR コンポーネントはすべての IMDS 呼び出しに IMDSv2 を使用します。アプリケーションコードでの IMDS 呼び出しでは、IMDSv1 と IMDSv2 の両方を使用するか、または、セキュリティを強化するために IMDSv2 のみを使用するように IMDS を設定できます。以前の Amazon EMR 6.x リリースで IMDSv1 を無効にすると、クラスターの起動が失敗します。



## 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Spark: Spark ランタイムのパフォーマンスが向上しました。

## 既知の問題

- Amazon EMR 6.2 の EMR 6.2.0 の `/etc/cron.d/libinstance-controller-java` file に正しくないアクセス許可が設定されている。ファイルに対するアクセス許可は、`644 (-rw-r--r--)` であるべきときに `645 (-rw-r--r-x)` になります。その結果、Amazon EMR バージョン 6.2 ではインスタンス状態ログが記録されず、`/emr/instance-logs` ディレクトリは空になります。この問題は、Amazon EMR 6.3.0 以降で修正されています。

この問題を回避するには、クラスター起動時に以下のスクリプトをブートストラップアクションとして実行します。

```
#!/bin/bash
sudo chmod 644 /etc/cron.d/libinstance-controller-java
```

- Amazon EMR 6.2.0 および 6.3.0 のプライベートサブネットクラスターで、Ganglia ウェブ UI にアクセスできません。「アクセス拒否 (403)」というエラーが表示されます。Spark、Hue、Zeppelin、Livy JupyterHub、Tez などの他のウェブ UIs は正常に動作しています。パブリックサブネットクラスターでの Ganglia ウェブ UI アクセスも正常に動作します。この問題を解決するには、`sudo systemctl restart httpd` を使用してプライマリノードで httpd サービスを再起動します。この問題は、Amazon EMR 6.4.0 で修正されています。
- Amazon EMR 6.2.0 で、httpd が連続して失敗し、Ganglia が利用できないという問題があります。「サーバーに接続できません」というエラーが表示されます。この問題が発生している、既に実行されているクラスターを修正するには、クラスターのプライマリノードに SSH 接続して、`/etc/httpd/conf/httpd.conf` にあるファイル `httpd.conf` に行 `Listen 80` を追加します。この問題は、Amazon EMR 6.3.0 で修正されています。
- セキュリティ設定を使用している場合、EMR 6.2.0 クラスターで HTTPD が失敗します。これにより、Ganglia ウェブアプリケーションのユーザーインターフェイスが使用できなくなります。Ganglia ウェブアプリケーションのユーザーインターフェイスにアクセスするには、クラスターのプライマリノード上の `/etc/httpd/conf/httpd.conf` ファイルに `Listen 80` を追加します。クラスターへの接続の詳細については、「[SSH を使用してプライマリノードに接続する](#)」を参照してください。

EMR Notebooks では、セキュリティ構成を使用する場合、EMR 6.2.0 クラスターとの接続の確立にも失敗します。ノートブックは、カーネルのリストと Spark ジョブの送信に失敗します。代わりに、別のバージョンの Amazon EMR で EMR Notebooks を使用することをお勧めします。

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース `emr-5.30.x`、`emr-5.31.0`、`emr-5.32.0`、`emr-6.0.0`、`emr-6.1.0`、および `emr-6.2.0` は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の `ulimit` 設定が低くなります。Amazon EMR リリース `5.30.1`、`5.30.2`、`5.31.1`、`5.32.1`、`6.0.1`、`6.1.1`、`6.2.1`、`5.33.0`、`6.3.0` 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大

オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
```

```
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

### ⚠ Important

Amazon EMR 6.1.0 および 6.2.0 には、Hudi の挿入、アップサート、および削除操作すべてに大きな影響を及ぼす可能性があるパフォーマンスの問題が含まれています。Amazon EMR 6.1.0 または 6.2.0 で Hudi を使用する予定がある場合は、AWS サポートに連絡してパッチが適用された Hudi RPM を取得する必要があります。

### ⚠ Important

Amazon Linux または Amazon Linux 2 Amazon マシンイメージ (AMI) を実行する EMR クラスタは、デフォルトの Amazon Linux 動作を使用します。再起動が必要な重要かつクリティカルなカーネル更新が自動的にダウンロードされてインストールされることはありません。これは、デフォルトの Amazon Linux AMI を実行している他の Amazon EC2 インスタンスと同じ動作です。Amazon EMR リリースが利用可能になった後に、再起動が必要な新しい Amazon Linux ソフトウェアアップデート (カーネル、NVIDIA、CUDA のアップデートなど) が使用可能になった場合、デフォルトの AMI を実行する EMR クラスタインスタンスで、それらの更新が自動的にダウンロードされてインストールされることはありません。カーネルの更新を取得するには、[Amazon EMR AMI をカスタマイズ](#)して、[最新の Amazon Linux AMI を使用](#)できるようにします。

- Amazon EMR 6.2.0 の Maven アーティファクトは公開されません。これらは Amazon EMR の将来のリリースで公開される予定です。
- HBase ストアファイルシステムテーブルを使用した永続的な HFile トラッキングは、HBase リージョンのレプリケーション機能をサポートしていません。HBase リージョンレプリケーションの詳細については、「[Timeline-consistent High Available Reads](#)」を参照してください。
- Amazon EMR 6.x と EMR 5.x の Hive バケット化バージョンの違い

EMR 5.x では OOS Apache Hive 2 を使用していますが、EMR 6.x では OOS Apache Hive 3 を使用しています。オープンソースの Hive2 ではバケット化バージョン 1 を使用していますが、オープンソースの Hive3 ではバケット化バージョン 2 を使用しています。Hive 2 (EMR 5.x) と Hive 3 (EMR 6.x) のこのバケット化バージョンの違いは、Hive のバケット化ハッシュ関数が異なることを意味します。以下の例を参照してください。

次の表は、それぞれ EMR 6.x および EMR 5.x で作成した例です。

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

EMR 6.x と EMR 5.x の両方に同じデータを挿入しています。

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

S3 の場所を確認すると、バケット化ファイル名が異なることが示されています。これは、ハッシュ関数が EMR 6.x (Hive 3) と EMR 5.x (Hive 2) で異なるためです。

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16          13 000025_0
2020-10-21 20:35:22          14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07          13 000066_0
2020-10-21 20:32:51          14 000072_0
```

EMR 6.x で、Hive CLI で次のコマンドを実行し、バージョンの違いを確認することもできます。バケット化バージョン 2 が返されることに注意してください。

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
  bucketing_version      2
...
```

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チ

チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: `!"#$%&'()*+,-`。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、`spark-defaults` 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## リリース 5.31.0

次のリリースノートには、Amazon EMR リリース 5.31.0 に関する情報が含まれています。5.30.1 からの変更が含まれています。

初回リリース日: 2020 年 10 月 9 日

最終更新日: 2020 年 10 月 15 日

### アップグレード

- Amazon Glue コネクタをバージョン 1.13.0 にアップグレードしました
- Amazon SageMaker Spark SDK をバージョン 1.4.0 にアップグレードしました
- Amazon Kinesis コネクタをバージョン 3.5.9 にアップグレードしました
- バージョン 1.11.852 AWS SDK for Java にアップグレード
- Bigtop-tomcat をバージョン 8.5.56 にアップグレードしました
- EMRFS をバージョン 2.43.0 にアップグレードしました
- EMR MetricsAndEventsApiGateway クライアントをバージョン 1.4.0 にアップグレードしました
- EMR S3 Dist CP をバージョン 2.15.0 にアップグレードしました
- EMR S3 Select をバージョン 1.6.0 にアップグレードしました
- Flink をバージョン 1.11.0 にアップグレードしました
- Hadoop をバージョン 2.10.0 にアップグレードしました
- Hive をバージョン 2.3.7 にアップグレードしました
- Hudi をバージョン 0.6.0 にアップグレードしました
- Hue をバージョン 4.7.1 にアップグレードしました
- バージョン 1.1.0 JupyterHub にアップグレード
- MXNet をバージョン 1.6.0 にアップグレードしました
- OpenCV をバージョン 4.3.0 にアップグレードしました
- Presto をバージョン 0.238.3 にアップグレードしました
- バージョン 2.1.0 TensorFlow にアップグレード

### 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。



- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- [Hive の列統計](#) は、Amazon EMR バージョン 5.31.0 以降でサポートされています。
- コンポーネントのバージョンをアップグレードしました。
- Amazon EMR 5.31.0 での EMRFS S3EC V2 のサポート。S3 Java SDK リリース 1.11.837 以降では、暗号化クライアントバージョン 2 (S3EC V2) がさまざまなセキュリティ機能強化とともに導入されました。詳細については、次を参照してください。
  - S3 ブログ投稿: [Updates to the Amazon S3 encryption client](#)。
  - AWS SDK for Java デベロッパーガイド: [暗号化クライアントと復号クライアントを V2 に移行します](#)。
  - EMR 管理ガイド: [Amazon S3 クライアント側暗号化](#)。

下位互換性を保つため、暗号化クライアント V1 は SDK で引き続き使用できます。



## 新機能

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
```

```
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Amazon EMR 5.31.0 を使用すると、Lake Formation と統合されるクラスターを起動できます。この統合により、AWS Glue Data Catalog のデータベースとテーブルにきめ細かな列レベルのデータフィルタリングが提供されます。また、これにより、企業の ID システムから EMR Notebooks または Apache Zeppelin へのフェデレーションシングルサインオンが可能になります。詳細については、「Amazon EMR 管理ガイド」の「[Amazon EMR と AWS Lake Formation の統合](#)」を参照してください。

Amazon EMR with Lake Formation は現在、米国東部 (オハイオ州とバージニア北部)、米国西部 (北カリフォルニアとオレゴン)、アジアパシフィック (ムンバイ、ソウル、シンガポール、シドニー、東京)、カナダ (中部)、欧州 (フランクフルト、アイルランド、ロンドン、パリ、ストックホルム)、南米 (サンパウロ) の 16 AWS リージョンで利用可能です。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Amazon EMR 5.31.0 AtRestEncryption または 5.32.0 を使用するクラスターで または HDFS 暗号化が有効になっている場合、Hive クエリでは次のランタイム例外が発生します。

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
java.lang.RuntimeException: Hive Runtime Error while closing
operators: java.io.IOException: java.util.ServiceConfigurationError:
org.apache.hadoop.security.token.TokenIdentifier: Provider
org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、`/` 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、`s3://bucket/table/p=a b` の `a` と `b` の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに

注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の spark.sql.sources.fastS3PartitionDiscovery.enabled 設定を false にセットします。

## リリース 6.1.0

次のリリースノートには、Amazon EMR リリース 6.1.0 に関する情報が含まれています。6.0.0 からの変更が含まれています。

初回リリース日: 2020 年 9 月 4 日

最終更新日: 2020 年 10 月 15 日

サポートされているアプリケーション

- AWS SDK for Java バージョン 1.11.828
- Flink バージョン 1.11.0
- Ganglia バージョン 3.7.2
- Hadoop バージョン 3.2.1-amzn-1
- HBase バージョン 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog バージョン 3.1.2-amzn-0
- Hive バージョン 3.1.2-amzn-1
- Hudi バージョン 0.5.2-incubating
- Hue バージョン 4.7.1
- JupyterHub バージョン 1.1.0
- Livy バージョン 0.7.0
- MXNet バージョン 1.6.0
- Oozie バージョン 5.2.0
- Phoenix バージョン 5.0.0
- Presto バージョン 0.232

- PrestoSQL バージョン 338
- Spark バージョン 3.0.0-amzn-0
- TensorFlow バージョン 2.1.0
- Zeppelin バージョン 0.9.0-preview1
- Zookeeper バージョン 3.4.14
- コネクタおよびドライバー: DynamoDB Connector 4.14.0

## 新機能

- ARM インスタンスタイプは、Amazon EMR バージョン 5.30.0 以降および Amazon EMR バージョン 6.1.0 以降でサポートされています。
- M6g 汎用インスタンスタイプは、Amazon EMR バージョン 6.1.0 以降および Amazon EMR バージョン 5.30.0 以降でサポートされています。詳細については、「Amazon EMR 管理ガイド」の「[サポートされるインスタンスタイプ](#)」を参照してください。
- EC2 プレイACEMENTグループ機能は、複数のプライマリノードクラスターのオプションとして Amazon EMR バージョン 5.23.0 以降でサポートされています。現在、プレイACEMENTグループ機能ではプライマリノードタイプのみがサポートされており、SPREAD ストラテジーは、これらのプライマリノードに適用されます。SPREAD 戦略では、ハードウェア障害の発生時に複数のプライマリノードが失われるのを防ぐため、少数のインスタンスを別個の基盤となるハードウェア全体に配置します。詳細については、「Amazon EMR 管理ガイド」の「[EC2 プレイACEMENTグループと EMR の統合](#)」を参照してください。
- マネージドスケーリング - Amazon EMR バージョン 6.1.0 では、Amazon EMR Managed Scaling を有効にすることで、ワークロードに応じてクラスター内のインスタンスやユニットの数を自動的に増減できます。Amazon EMR は引き続きクラスターのメトリクスを評価し、クラスターのコストと速度を最適化するためのスケーリングを決定します。マネージドスケーリングは、Amazon EMR バージョン 5.30.0 以降 (6.0.0 を除く) でも使用できます。詳細については、「Amazon EMR 管理ガイド」の「[クラスターリソースのスケーリング](#)」を参照してください。
- PrestoSQL バージョン 338 は EMR 6.1.0 でサポートされています。詳細については、「[Presto](#)」を参照してください。
  - PrestoSQL は EMR 6.1.0 以降のバージョンでのみサポートされています。EMR 6.0.0 および EMR 5.x ではサポートされていません。
  - アプリケーション名 Presto は、クラスターに PrestoDB をインストールするために引き続き使用されます。クラスターに PrestoDB をインストールするには、アプリケーション名 PrestoSQL を使用します。

- PrestoDB または PrestoSQL のいずれかをインストールできますが、両方を 1 つのクラスターにインストールすることはできません。クラスターの作成時に PrestoDB と PrestoSQL の両方を指定すると、検証エラーが発生し、クラスター作成リクエストは失敗します。
- PrestoSQL は、シングルマスタークラスターとマルチマスタークラスターの両方でサポートされています。マルチマスタークラスターでは、PrestoSQL または PrestoDB を実行するには、外部 Hive メタストアが必要です。「[複数のプライマリノードを持つ Amazon EMR クラスターでサポートされるアプリケーション](#)」を参照してください。
- Docker を使用した Apache Hadoop および Apache Spark での ECR 自動認証サポート: Spark ユーザーは、Docker Hub および Amazon Elastic Container Registry (Amazon ECR) の Docker イメージを使用して、環境とライブラリの依存関係を定義できます。

[Docker を設定し、Amazon EMR 6.x を使用して Docker で Spark アプリケーションを実行します。](#)

- EMR は Apache Hive ACID トランザクションをサポートしています。Amazon EMR 6.1.0 では、データベースの ACID プロパティに準拠するように Hive ACID トランザクションのサポートが追加されています。この機能を使用すると、Amazon Simple Storage Service (Amazon S3) のデータを使用して Hive マネージドテーブルで INSERT, UPDATE, DELETE, および MERGE の各オペレーションを実行できます。これは、ストリーミングの取り込み、データの再記述、MERGE を使用した一括更新、徐々に変化するディメンションなどのユースケースにとって重要な機能です。設定例とユースケースを含め、詳細については、「[Amazon EMR supports Apache Hive ACID transactions](#)」を参照してください。

## 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。
- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールアップリクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールアップ操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。



- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールアップ中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Apache Flink は EMR 6.0.0 ではサポートされていませんが、Flink 1.11.0 と EMR 6.1.0 ではサポートされています。これは、Hadoop 3 を公式にサポートする Flink の最初のバージョンです。「[Apache Flink 1.11.0 Release Announcement](#)」を参照してください。
- Ganglia はデフォルトの EMR 6.1.0 パッケージバンドルから削除されました。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

## コマンドラインから ulimit を明示的に設定する

1. `/etc/systemd/system/instance-controller.service` を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

## ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```



**⚠ Important**

Amazon EMR 6.1.0 および 6.2.0 には、Hudi の挿入、アップサート、および削除操作すべてに大きな影響を及ぼす可能性があるパフォーマンスの問題が含まれています。Amazon EMR 6.1.0 または 6.2.0 で Hudi を使用する予定がある場合は、AWS サポートに連絡してパッチが適用された Hudi RPM を取得する必要があります。

- `spark.driver.extraJavaOptions` と `spark.executor.extraJavaOptions` を使用してカスタムのガベージコレクション設定を指定すると、ガベージコレクション設定の競合のために、EMR 6.1 でドライバー/エグゼキュターの起動に失敗します。EMR リリース 6.1.0 では、代わりにプロパティ `spark.driver.defaultJavaOptions` と `spark.executor.defaultJavaOptions` を使用して、ドライバーとエグゼキュターのカスタムの Spark ガベージコレクション設定を指定する必要があります。詳細については、「[Apache Spark のランタイム環境](#)」および「[Amazon EMR 6.1.0 での Spark ガベージコレクションの設定](#)」を参照してください。
- Oozie で Pig を (Hue 内で。Hue では、Oozie アクションを使用して Pig スクリプトを実行するため) 使用すると、`native-izo` ライブラリをロードできないというエラーが生成されます。このエラーメッセージは情報であるため、Pig の実行はブロックされません。
- Hudi 同時実行のサポート: 現在、Hudi は単一の Hudi テーブルへの同時書き込みをサポートしていません。さらに、Hudi は新しいライターの起動を許可する前に、進行中のライターによって行われた変更をロールバックします。同時書き込みは、このメカニズムを妨げて、競合状態を引き起こし、データの破損につながる可能性があります。データ処理ワークフローの一部として、いつでも 1 つの Hudi テーブルに対して動作する Hudi ライターが 1 つだけであることを確認してください。Hudi は、同じ Hudi テーブルに対して動作する複数の同時リーダーをサポートしています。
- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

**回避方法:**

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。

- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Amazon EMR 6.1.0 には、Presto を実行しているクラスターに影響する問題があります。長時間 (日単位) 後に、クラスターが「su: failed to execute /bin/bash: Resource temporarily unavailable」(su: /bin/bash の実行に失敗しました:リソースを一時的に利用できません) または「shell request failed on channel 0」(チャンネル 0 でシェルリクエストが失敗しました) などのエラーをスローする可能性があります。この問題は、内部の Amazon EMR プロセス (InstanceController) が生成するライトウェイトプロセス (LWP) が多すぎるため、最終的に Hadoop ユーザーが nproc 制限を超える原因となります。これにより、ユーザーは追加のプロセスを開くことができなくなります。この問題の解決策は、EMR 6.2.0 にアップグレードすることです。

## リリース 6.0.0

次のリリースノートには、Amazon EMR リリース 6.0.0 に関する情報が含まれています。

初回リリース日: 2020 年 3 月 10 日

サポートされているアプリケーション

- AWS SDK for Java バージョン 1.11.711
- Ganglia バージョン 3.7.2
- Hadoop バージョン 3.2.1
- HBase バージョン 2.2.3
- HCatalog バージョン 3.1.2
- Hive バージョン 3.1.2

- Hudi バージョン 0.5.0-incubating
- Hue バージョン 4.4.0
- JupyterHub バージョン 1.0.0
- Livy バージョン 0.6.0
- MXNet バージョン 1.5.1
- Oozie バージョン 5.1.0
- Phoenix バージョン 5.0.0
- Presto バージョン 0.230
- Spark バージョン 2.4.4
- TensorFlow バージョン 1.14.0
- Zeppelin バージョン 0.9.0-SNAPSHOT
- Zookeeper バージョン 3.4.14
- コネクタおよびドライバー: DynamoDB Connector 4.14.0

#### Note

Flink、Sqoop、Pig、および Mahout は、Amazon EMR バージョン 6.0.0 では使用できません。

## 新機能

- YARN Docker ランタイムのサポート - Spark ジョブなどの YARN アプリケーションは、Docker コンテナのコンテキストで実行できるようになりました。これにより、Amazon EMR クラスターにカスタムライブラリをインストールすることなく、Docker イメージの依存関係を簡単に定義できます。詳細については、「[Docker 統合の設定](#)」および「[Amazon EMR 6.0.0 を使用して Docker で Spark アプリケーションを実行する](#)」を参照してください。
- Hive LLAP のサポート - クエリのパフォーマンス向上のため、Hive が LLAP 実行モードをサポートしました。詳細については、「[Hive LLAP の使用](#)」を参照してください。

## 変更点、機能強化、解決した問題

- これは、Amazon EMR Scaling がクラスターを正常にスケールアップ/スケールダウンできない場合や、アプリケーション障害を引き起こした場合の問題点を修正するためのリリースです。

- Amazon EMR のクラスター上のデーモンが YARN ノード状態や HDFS ノード状態の収集などのヘルスチェックアクティビティを実行しているときに、大規模で使用率の高いクラスターのスケールリングクエストが失敗する問題を修正しました。これは、クラスター上のデーモンがノードのヘルスステータスデータを内部の Amazon EMR コンポーネントに伝達できなかったために発生していました。
- EMR のクラスター上のデーモンが改善され、IP アドレスが再利用されるときにノードの状態を正しく追跡できるようになり、スケールリング操作中の信頼性が向上しました。
- [SPARK-29683](#)。Spark が使用可能なすべてのノードが拒否リストに登録されていると想定していたため、クラスターのスケールダウン中にジョブエラーが発生する問題を修正しました。
- [YARN-9011](#)。クラスターがスケールアップまたはスケールダウンを試みたときに YARN 廃止の競合状態が原因でジョブ障害が発生する問題を修正しました。
- Amazon EMR のクラスター上のデーモンと YARN/HDFS の間でノードの状態が常に一致するようにすることで、クラスターのスケールリング中のステップまたはジョブの障害に関する問題を修正しました。
- Kerberos 認証で有効になっている Amazon EMR クラスターで、スケールダウンやステップ送信などのクラスター操作が失敗する問題を修正しました。これは、Amazon EMR のクラスター上のデーモンが、プライマリノードで実行されている HDFS/YARN と安全に通信するために必要な Kerberos チケットを更新しなかったためです。
- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- Amazon Linux
  - Amazon Linux 2 は EMR 6.x リリースシリーズのオペレーティングシステムです。
  - systemd は、Amazon Linux 1 で使用される upstart ではなく、サービス管理に使用されます。
- Java Development Kit (JDK)
  - Corretto JDK 8 は、EMR 6.x リリースシリーズのデフォルトの JDK です。
- Scala
  - Scala 2.12 は、Apache Spark および Apache Livy で使用されます。
- Python 3
  - Python 3 が EMR の Python のデフォルトバージョンになりました。
- YARN ノードラベル

- Amazon EMR 6.x リリースシリーズ以降では、YARN ノードラベル機能はデフォルトで無効になっています。アプリケーションマスタープロセスは、デフォルトでコアノードとタスクノードの両方で実行できます。次のプロパティを設定することで、YARN ノードラベル機能を有効にできます: `yarn.node-labels.enabled` および `yarn.node-labels.am.default-node-label-expression`。詳細については、「[Understanding Primary, Core, and Task Nodes](#)」を参照してください。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース `emr-5.30.x`、`emr-5.31.0`、`emr-5.32.0`、`emr-6.0.0`、`emr-6.1.0`、および `emr-6.2.0` は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の `ulimit` 設定が低くなります。Amazon EMR リリース `5.30.1`、`5.30.2`、`5.31.1`、`5.32.1`、`6.0.1`、`6.1.1`、`6.2.1`、`5.33.0`、`6.3.0` 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの `ulimit` 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の `ulimit` 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に `ulimit` 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、`instance-controller ulimit` を最大の 65536 ファイルに明示的に設定できます。

### コマンドラインから `ulimit` を明示的に設定する

1. `/etc/systemd/system/instance-controller.service` を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- 、SparkR PySpark、Spark-shell などの Spark インタラクティブシェルは、追加のライブラリでの Docker の使用をサポートしていません。SparkR
- Amazon EMR バージョン 6.0.0 で Python 3 を使用するには、PATH を `yarn.nodemanager.env-whitelist` に追加する必要があります。
- AWS Glue Data Catalog を Hive のメタストアとして使用する場合、Live Long and Process (LLAP) 機能はサポートされません。
- Spark と Docker の統合で Amazon EMR 6.0.0 を使用する場合、Docker ランタイムで Spark ジョブを送信する際の失敗を回避するために、同じインスタンスタイプおよび同量の EBS ボリュームでクラスター内のインスタンスを設定する必要があります。
- Amazon EMR 6.0.0 では、HBase on Amazon S3 ストレージモードは [HBASE-24286](#) の問題の影響を受けます。既存の S3 データを使用してクラスターを作成する場合、HBase マスターは初期化できません。

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.30.1

次のリリースノートには、Amazon EMR リリース 5.30.1 に関する情報が含まれています。5.30.0 からの変更が含まれています。

初回リリース日: 2020 年 6 月 30 日

最終更新日: 2020 年 8 月 24 日

変更点、機能強化、解決した問題

- Amazon EMR の新しいリリースでは、Amazon EMR の古い AL2 で「最大オープンファイル」の上限が低い問題が修正されています。Amazon EMR リリース



- 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれるようになりました。
- インスタンスコントローラープロセスが無限数のプロセスをスポンする問題を修正しました。
  - Hue が Hive クエリを実行できず、「データベースがロックされています」というメッセージが表示され、クエリの実行が妨げられる問題を修正しました。
  - EMR クラスターでより多くのタスクを同時に実行できるようにするための Spark の問題を修正しました。
  - Jupyter サーバーで「too many files open error」(開いているファイルの数が多すぎるエラー)が発生する、Jupyter Notebook の問題を修正しました。
  - クラスターの開始時間の問題を修正しました。

## 新機能

- Tez UI および YARN タイムラインサーバーの永続アプリケーションインターフェイスは、Amazon EMR バージョン 6.x および EMR バージョン 5.30.1 以降で利用可能です。永続アプリケーション履歴へのワンクリックリンクアクセスにより、SSH 接続でウェブプロキシを設定せずに、ジョブ履歴にすばやくアクセスできます。アクティブなクラスターと終了したクラスターのログは、アプリケーションの終了後 30 日間利用できます。詳細については、「Amazon EMR 管理ガイド」の「[永続アプリケーションユーザーインターフェイスの表示](#)」を参照してください。
- EMR Notebooks 実行 API を使用して、スクリプトまたはコマンドラインで EMR Notebooks を実行できます。AWS コンソールなしで EMR Notebooks の実行を開始、停止、一覧表示、および記述する機能により、EMR Notebooks をプログラムで制御できます。パラメータ化されたノートブックセルを使用すると、新しいパラメータ値セットごとにノートブックのコピーを作成しなくても、複数の異なるパラメータ値をノートブックに渡すことができます。「[EMR API のアクション](#)」を参照してください。サンプルコードについては、「[EMR Notebooks をプログラムで実行するサンプルコマンド](#)」を参照してください。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース emr-5.30.x、emr-5.31.0、emr-5.32.0、emr-6.0.0、emr-6.1.0、および emr-6.2.0 は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の ulimit 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オーブ



ンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの ulimit 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の ulimit 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に ulimit 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、instance-controller ulimit を最大の 65536 ファイルに明示的に設定できます。

コマンドラインから ulimit を明示的に設定する

1. /etc/systemd/system/instance-controller.service を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
```

```
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

## • EMR Notebooks

クラスターのプライマリノードにカーネルと追加の Python ライブラリをインストールする機能は、EMR バージョン 5.30.1 ではデフォルトで無効になっています。この機能の詳細については、「[クラスターのプライマリノードへのカーネルと Python ライブラリのインストール](#)」を参照してください。

この機能を有効にするには、以下の操作を行います。

1. EMR Notebooks のサービスロールにアタッチされているアクセス許可ポリシーで、次のアクションが許可されていることを確認します。

```
elasticmapreduce:ListSteps
```

詳細については、「[EMR Notebooks のサービスロール](#)」を参照してください。

2. 次の例に示すように、を使用して EMR Notebooks をセットアップするステップをクラスターで AWS CLI 実行します。`us-east-1` を、クラスターが存在するリージョンに置き換えます。詳細については、「[AWS CLIを使用したクラスターへのステップの追加](#)」を参照してください。

```
aws emr add-steps --cluster-id MyClusterID --steps
  Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
```

## • マネージドスケールリング

Presto がインストールされていない 5.30.0 クラスターおよび 5.30.1 クラスターでマネージドスケールリング操作を実行すると、特にスケールダウン操作の後にすぐ、スケールアップ操作が実行されたときに、アプリケーション障害が発生するか、ユニフォームインスタンスグループまたはインスタンスフリートが ARRESTED 状態のままになる場合があります。

回避策として、ジョブに Presto が 必要ない場合でも、Amazon EMR リリース 5.30.0 および 5.30.1 を使用してクラスターを作成するときに、インストールするアプリケーションとして Presto を選択します。

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、`s3://bucket/table/p=a` は `s3://bucket/table/p=a b` のプレフィックスです。

- 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## リリース 5.30.0

次のリリースノートには、Amazon EMR リリース 5.30.0 に関する情報が含まれています。5.29.0 からの変更が含まれています。

初回リリース日: 2020 年 5 月 13 日

最終更新日: 2020 年 6 月 25 日

### アップグレード

- バージョン 1.11.759 AWS SDK for Java にアップグレード
- Amazon SageMaker Spark SDK をバージョン 1.3.0 にアップグレードしました
- EMR Record Server をバージョン 1.6.0 にアップグレードしました
- Flink をバージョン 1.10.0 にアップグレードしました
- Ganglia をバージョン 3.7.2 にアップグレードしました
- HBase をバージョン 1.4.13 にアップグレードしました
- Hudi をバージョン 0.5.2 (incubating) にアップグレードしました
- Hue をバージョン 4.6.0 にアップグレードしました
- バージョン 1.1.0 JupyterHub にアップグレード
- Livy をバージョン 0.7.0 (incubating) にアップグレードしました
- Oozie をバージョン 5.2.0 にアップグレードしました
- Presto をバージョン 0.232 にアップグレードしました
- Spark をバージョン 2.4.5 にアップグレードしました

- コネクタとドライバーをアップグレードしました: Amazon Glue Connector 1.12.0、Amazon Kinesis Connector 3.5.0、EMR DynamoDB Connector 4.14.0

## 新機能

- EMR Notebooks - 5.30.0 を使用して作成された EMR クラスターで使用する場合、EMR ノートブックカーネルはクラスター上で実行されます。これにより、ノートブックのパフォーマンスが向上し、カーネルをインストールおよびカスタマイズすることができます。また、クラスターのプライマリノードに Python ライブラリをインストールすることもできます。詳細については、「EMR 管理ガイド」の「[カーネルとライブラリのインストールと使用](#)」を参照してください。
- マネージドスケールリング - Amazon EMR バージョン 5.30.0 以降では、EMR マネージドスケールリングを有効にすることで、ワークロードに応じてクラスター内のインスタンスやユニットの数を自動的に増減できます。Amazon EMR は引き続きクラスターのメトリクスを評価し、クラスターのコストと速度を最適化するためのスケールリングを決定します。詳細については、「Amazon EMR 管理ガイド」の「[クラスターリソースのスケールリング](#)」を参照してください。
- Amazon S3 に保存されているログファイルの暗号化 - Amazon EMR バージョン 5.30.0 以降では、Amazon S3 に保存されているログファイルを AWS KMS カスタマーマネージドキーで暗号化できます。詳細については、「Amazon EMR 管理ガイド」の「[Amazon S3 に保存されているログファイルの暗号化](#)」を参照してください。
- Amazon Linux 2 のサポート - EMR バージョン 5.30.0 以降では、EMR は Amazon Linux 2 OS を使用します。新しいカスタム AMI (Amazon マシンイメージ) は、Amazon Linux 2 AMI に基づいている必要があります。詳細については、「[カスタム AMI の使用](#)」を参照してください。
- Presto Graceful Auto Scale - 5.30.0 を使用する EMR クラスターでは、オートスケールリングのタイムアウト期間を設定することで、Presto タスクの実行が終了するまで待つからノードの使用を停止できます。詳細については、「[グレースフルな廃止による Presto Auto Scaling の使用](#)」を参照してください。
- 新しい割り当て戦略オプションを使用したフリートインスタンスの作成 — EMR バージョン 5.12.1 以降では、新しい割り当て戦略オプションを使用できます。これにより、クラスターのプロビジョニングにかかる時間が短縮され、スポット割り当てがより正確になり、スポットインスタンスの中断が低減します。デフォルト以外の EMR サービスロールを更新する必要があります。「[インスタンスフリートを構成する](#)」を参照してください。
- `sudo systemctl stop` コマンドと `sudo systemctl start` コマンド — Amazon Linux 2 OS を使用する EMR バージョン 5.30.0 以降では、EMR は `sudo systemctl stop` コマンドおよび `sudo systemctl start` コマンドを使用してサービスを再起動します。詳細については、「[Amazon EMR でサービスを再起動するにはどうすればよいですか?](#)」を参照してください。

## 変更、機能強化、解決した問題

- EMR バージョン 5.30.0 では、デフォルトで Ganglia がインストールされません。クラスターの作成時に、Ganglia を明示的に選択してインストールできます。
- Spark パフォーマンスの最適化。
- Presto パフォーマンスの最適化。
- Python 3 は、Amazon EMR バージョン 5.30.0 以降ではデフォルトです。
- プライベートサブネット内のサービスアクセス用のデフォルトのマネージドセキュリティグループが更新され、複数の新しいルールが追加されました。サービスアクセスにカスタムセキュリティグループを使用している場合は、同じルールをデフォルトのマネージドセキュリティグループとして含める必要があります。詳細については、「[サービスアクセスの Amazon EMR マネージドセキュリティグループ \(プライベートサブネット\)](#)」を参照してください。Amazon EMR でカスタムサービスロールを使用する場合は、`ec2:describeSecurityGroups` にアクセス許可を付与して、セキュリティグループが正常に作成されたかどうかを EMR で検証できるようにする必要があります。EMR\_DefaultRole を使用する場合は、このアクセス許可はデフォルトのマネージドポリシーに既に含まれています。

## 既知の問題

- 古い AL2 で「最大オープンファイル」の上限が低い [新しいリリースで修正済み]。Amazon EMR リリース `emr-5.30.x`、`emr-5.31.0`、`emr-5.32.0`、`emr-6.0.0`、`emr-6.1.0`、および `emr-6.2.0` は、古いバージョンの Amazon Linux 2 (AL2) に基づいており、デフォルトの AMI を使用して Amazon EMR クラスターを作成する場合に「最大オープンファイル」の `ulimit` 設定が低くなります。Amazon EMR リリース 5.30.1、5.30.2、5.31.1、5.32.1、6.0.1、6.1.1、6.2.1、5.33.0、6.3.0 以降には、「最大オープンファイル」設定が高くなった永続的な修正が含まれています。オープンファイルの上限が低いリリースでは、Spark ジョブを送信するときに「Too many open files」というエラーが発生します。影響を受けるリリースでは、Amazon EMR のデフォルト AMI の「最大オープンファイル」はデフォルトの `ulimit` 設定 4096 になっており、最新の Amazon Linux 2 AMI の上限 65536 ファイルよりも低くなっています。「最大オープンファイル」の `ulimit` 設定が低い場合、Spark ドライバーとエグゼキュータが 4096 を超えるファイルを開こうとすると、Spark ジョブが失敗します。この問題を解決するために、Amazon EMR には、クラスターの作成時に `ulimit` 設定を調整するブートストラップアクション (BA) スクリプトが用意されています。

この問題の永続的な修正がない古い Amazon EMR バージョンを使用している場合は、以下の回避策を使用すると、`instance-controller ulimit` を最大の 65536 ファイルに明示的に設定できます。

## コマンドラインから ulimit を明示的に設定する

1. `/etc/systemd/system/instance-controller.service` を編集して、Service セクションに次のパラメータを追加します。

```
LimitNOFILE=65536
```

```
LimitNPROC=65536
```

2. 再起動 InstanceController

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

## ブートストラップアクション (BA) を使用して ulimit を設定する

ブートストラップアクション (BA) スクリプトを使用して、クラスター作成時にインスタンスコントローラーの ulimit を 65536 ファイルに設定することもできます。

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- マネージドスケールリング



Presto がインストールされていない 5.30.0 クラスターおよび 5.30.1 クラスターでマネージドスケールリング操作を実行すると、特にスケールダウン操作の後にすぐ、スケールアップ操作が実行されたときに、アプリケーション障害が発生するか、ユニフォームインスタンスグループまたはインスタンスフリートが ARRESTED 状態のままになる場合があります。

回避策として、ジョブに Presto が必要ない場合でも、Amazon EMR リリース 5.30.0 および 5.30.1 を使用してクラスターを作成するときに、インストールするアプリケーションとして Presto を選択します。

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

- Hue 4.6.0 のデフォルトのデータベースエンジンは SQLite であり、外部データベースで Hue を使用しようとする、SQLite が原因で問題が発生します。これを修正するには、`hue-ini` 設定分類の `engine` を `mysql` に設定します。この問題は、Amazon EMR バージョン 5.30.1 で修正されています。



- Hive パーティション場所の形式設定で Spark を使用して Amazon S3 のデータを読み取り、Amazon EMR リリース 5.30.0 から 5.36.0、および 6.2.0 から 6.9.0 で Spark を実行すると、クラスターがデータを正しく読み取れなくなる問題が発生する可能性があります。これは、パーティションに以下の特徴がすべて当てはまる場合に発生する可能性があります。
  - 同じテーブルから 2 つ以上のパーティションがスキャンされます。
  - 少なくとも 1 つのパーティションディレクトリパスが、少なくとも 1 つの他のパーティションディレクトリパスのプレフィックスです。例えば、s3://bucket/table/p=a は s3://bucket/table/p=a b のプレフィックスです。
  - 他のパーティションディレクトリのプレフィックスに続く最初の文字が、/ 文字 (U+002F) より小さい UTF-8 値を持ちます。例えば、s3://bucket/table/p=a b の a と b の間にあるスペース文字 (U+0020) はこのカテゴリに該当します。非制御文字は他にも 14 個あることに注意してください: !"#\$%&'()\*+,-。詳細については、「[UTF-8 encoding table and Unicode characters](#)」を参照してください。

この問題の回避策として、spark-defaults 分類の `spark.sql.sources.fastS3PartitionDiscovery.enabled` 設定を `false` にセットします。

## リリース 5.29.0

次のリリースノートには、Amazon EMR リリース 5.29.0 に関する情報が含まれています。5.28.1 からの変更が含まれています。

初回リリース日: 2020 年 1 月 17 日

### アップグレード

- バージョン 1.11.682 AWS SDK for Java にアップグレード
- Hive をバージョン 2.3.6 にアップグレードしました
- Flink をバージョン 1.9.1 にアップグレードしました
- EMRFS をバージョン 2.38.0 にアップグレードしました
- EMR DynamoDB Connector をバージョン 4.13.0 にアップグレードしました

### 変更、機能強化、解決した問題

- Spark

- Spark パフォーマンスの最適化。
- EMRFS
  - 管理ガイドでは、整合性のあるビューを実現するために `emrfs-site.xml` デフォルト設定が更新されています。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.28.1

次のリリースノートには、Amazon EMR リリース 5.28.1 に関する情報が含まれています。5.28.0 からの変更が含まれています。

初回リリース日: 2020 年 1 月 10 日

## 変更、機能強化、解決した問題

- Spark
  - Spark の互換性の問題が修正されました。
- CloudWatch メトリクス
  - 複数のプライマリノードを持つ EMR クラスターでの Amazon CloudWatch メトリクスの発行を修正しました。
- 無効化されたログメッセージ
  - 「...using old version (<4.5.8) of Apache http client」 (Apache http クライアントの古いバージョン (<4.5.8) を使用) という誤ったログメッセージを無効にしました。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

**Note**

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.28.0

次のリリースノートには、Amazon EMR リリース 5.28.0 に関する情報が含まれています。変更は 5.27.0 に関連するものです。

初回リリース日: 2019 年 11 月 12 日

### アップグレード

- Flink をバージョン 1.9.0 にアップグレードしました
- Hive をバージョン 2.3.6 にアップグレードしました
- MXNet をバージョン 1.5.1 にアップグレードしました
- Phoenix をバージョン 4.14.3 にアップグレードしました
- Presto をバージョン 0.227 にアップグレードしました
- Zeppelin をバージョン 0.8.2 にアップグレードしました

### 新機能

- クラスターを作成するときに、Amazon EMR で [Apache Hudi](#) をインストールできるようになりました。詳細については、「[Hudi](#)」を参照してください。
- (2019 年 11 月 25 日) 複数のステップを並行して選択して、クラスター使用率を改善し、コストを削減できるようになりました。また、保留中および実行中のステップの両方をキャンセルできるようになりました。詳細については、「[AWS CLI およびコンソールを使用したステップの操作](#)」を参照してください。
- (2019 年 12 月 3 日) EMR クラスターを で作成して実行できるようになりました AWS Outposts。は、オンプレミス施設でネイティブ AWS サービス、インフラストラクチャ、運用モデル AWS Outposts を有効にします。AWS Outposts 環境では、AWS クラウドで使用するのと同じ AWS APIs ツール、インフラストラクチャを使用できます。詳細については、の「[EMR クラスター AWS Outposts](#)」を参照してください。

- (2020 年 3 月 11 日) Amazon EMR バージョン 5.28.0 以降では、AWS Local Zones をサポートする AWS リージョンの論理拡張として、Local Zones サブネットで Amazon EMR クラスターを作成して実行できます。Local Zone を使用すると、Amazon EMR の機能や、コンピューティングやストレージ AWS サービスなどのサービスのサブセットをユーザーの近くに配置することができ、ローカルで実行されているアプリケーションへのレイテンシーが非常に低くなります。使用可能な Local Zones のリストについては、「[AWS Local Zones](#)」を参照してください。利用可能な AWS ローカルゾーンへのアクセスの詳細については、「[リージョン、アベイラビリティゾーン、およびローカルゾーン](#)」を参照してください。

Local Zones は、現在 Amazon EMR Notebooks をサポートしておらず、インターフェイス VPC エンドポイント (AWS PrivateLink) を使用した Amazon EMR への直接接続をサポートしていません。

### 変更、機能強化、解決した問題

- 高可用性クラスターに対するアプリケーションのサポートを拡大しました。
  - 詳細については、「Amazon EMR 管理ガイド」の「[複数のプライマリノードを持つ Amazon EMR クラスターでサポートされるアプリケーション](#)」を参照してください。
- Spark
  - パフォーマンスの最適化
- [Hive]
  - パフォーマンスの最適化
- Presto
  - パフォーマンスの最適化

### 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.27.0

次のリリースノートには、Amazon EMR リリース 5.27.0 に関する情報が含まれています。5.26.0 からの変更が含まれています。

初回リリース日: 2019 年 9 月 23 日

### アップグレード

- AWS SDK for Java 1.11.615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- Tensorflow 1.14.0
- コネクタおよびドライバー:
  - DynamoDB Connector 4.12.0

## 新機能

- (2019 年 10 月 24 日) EMR ノートブックの次の新機能は、すべての Amazon EMR リリースで利用できます。
  - Git リポジトリを EMR ノートブックに関連付けて、バージョン管理された環境でノートブックを保存できます。リモートの Git リポジトリを使用して、コードをピアと共有し、既存の Jupyter Notebook を再利用できます。詳細については、「Amazon EMR 管理ガイド」の「[Git リポジトリを Amazon EMR Notebooks に関連付ける](#)」を参照してください。
  - ノートブックの比較とマージを簡素化するために、[nbdime ユーティリティ](#)が EMR ノートブックで利用できるようになりました。
  - EMR ノートブックでサポートされるようになりました JupyterLab。JupyterLab は、Jupyter Notebook と完全に互換性のあるウェブベースのインタラクティブな開発環境です。これで、ノートブックを JupyterLab または Jupyter Notebook Editor で開くことが選択できるようになりました。
- (2019 年 10 月 30 日) Amazon EMR バージョン 5.25.0 以降では、コンソールでクラスターの [概要] ページまたは [アプリケーションの履歴] タブから Spark 履歴サーバー UI に接続できます。SSH 接続でウェブプロキシを設定する代わりに、Spark 履歴サーバー UI にすばやくアクセスしてアプリケーションのメトリクスを確認し、アクティブなクラスターと終了したクラスターに関連するログファイルにアクセスできます。詳細については、「Amazon EMR 管理ガイド」の「[永続アプリケーションユーザーインターフェイスへのクラスター外アクセス](#)」を参照してください。

## 変更点、機能強化、解決した問題

- 複数のプライマリノードを持つ Amazon EMR クラスター
  - 複数のプライマリノードを持つ Amazon EMR クラスターに Flink をインストールして実行できます。詳細については、「[サポートされるアプリケーションと機能](#)」を参照してください。
  - 複数のプライマリノードを持つ Amazon EMR クラスターに HDFS 透過的暗号化を設定することができます。詳細については、「[HDFS Transparent Encryption on EMR clusters with Multiple Primary Nodes](#)」を参照してください。
  - 複数のプライマリノードを持つ Amazon EMR クラスターで実行されているアプリケーションの設定を変更できるようになりました。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。
- Amazon EMR-DynamoDB Connector

- Amazon EMR-DynamoDB Connector は、boolean、list、map、item、null の DynamoDB データ型をサポートするようになりました。詳細については、「[Hive コマンドを実行するように Hive テーブルをセットアップする](#)」を参照してください。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.26.0

次のリリースノートには、Amazon EMR リリース 5.26.0 に関する情報が含まれています。5.25.0 からの変更が含まれています。



初回リリース日: 2019 年 8 月 8 日

最終更新日: 2019 年 8 月 19 日

## アップグレード

- AWS SDK for Java 1.11.595
- HBase 1.4.10
- Phoenix 4.14.2
- コネクタおよびドライバー:
  - DynamoDB Connector 4.11.0
  - MariaDB Connector 2.4.2
  - Amazon Redshift JDBC ドライバー 1.2.32.1056

## 新機能

- (ベータ) Amazon EMR 5.26.0 を使用すると、Lake Formation と統合されるクラスターを起動できます。この統合により、AWS Glue Data Catalog のデータベースとテーブルへのきめ細かな列レベルのアクセスが可能になります。また、これにより、企業の ID システムから EMR Notebooks または Apache Zeppelin へのフェデレーションシングルサインオンが可能になります。詳細については、[「Amazon EMR と AWS Lake Formation \(ベータ\) の統合」](#)を参照してください。
- (2019 年 8 月 19 日) Amazon EMR のパブリックアクセスブロックが、セキュリティグループをサポートしているすべての Amazon EMR リリースで利用できるようになりました。ブロックパブリックアクセスは、各 AWS リージョンに適用されるアカウント全体の設定です。パブリックアクセスブロックにより、クラスターに関連付けられているセキュリティグループに、ポートで IPv4 0.0.0.0/0 または IPv6 ::/0 (パブリックアクセス) からのインバウンドトラフィックを許可するルールがある場合に、クラスターの起動が禁止されます (ポートが例外として指定されている場合を除く)。ポート 22 は、デフォルトで例外になります。詳細については、「Amazon EMR 管理ガイド」の [「Amazon EMR のパブリックアクセスブロックの使用」](#)を参照してください。

## 変更、機能強化、解決した問題

- EMR Notebooks
  - EMR 5.26.0 以降では、EMR Notebooks はデフォルトの Python ライブラリに加えて、ノートブックスコープの Python ライブラリもサポートしています。ノートブックスコープのライブラリは、クラスターを再作成したりノートブックをクラスターに再アタッチしたりすることなく、

ノートブックエディター内からインストールできます。ノートブックスコープのライブラリは Python 仮想環境に作成されるため、現在のノートブックセッションにのみ適用されます。これにより、ノートブックの依存関係を分離できます。詳細については、「Amazon EMR 管理ガイド」の「[ノートブックスコープのライブラリの使用](#)」を参照してください。

- EMRFS

- `fs.s3.consistent.metadata.etag verification.enabled` を `true` に設定すると、ETag 検証機能 (ベータ) を有効にできます。この機能を有効にすると、EMRFS は Amazon S3 ETags を使用して、読み取られているオブジェクトが利用可能な最新バージョンであることを確認します。この機能は、Amazon S3 上のファイルが同じ名前を保持したまま上書きされる read-after-update コースケースに役立ちます。この ETag 検証機能は、現在 S3 Select では使用できません。詳細については、「[整合性のあるビューを設定する](#)」を参照してください。

- Spark

- 現在、次の最適化がデフォルトで有効になります: ダイナミックパーティションプルーニング、DISTINCT before INTERSECT、DISTINCT クエリが後に続く JOIN の SQL 計画統計推論の改善、スカラーサブクエリのフラット化、結合順序の最適化、ブルームフィルター結合。詳細については、「[Spark のパフォーマンスの最適化](#)」を参照してください。
  - ソートマージ結合の全体的なステージコード生成が改善されました。
  - クエリフラグメントとサブクエリの再利用が改善されました。
  - Spark の起動時にエグゼキュターを事前割り当てする機能が改善されました。
  - 結合の小さい側にブロードキャストヒントが含まれている場合、ブルームフィルター結合は適用されなくなりました。
- Tez
  - Tez に関する問題を解決しました。Tez UI は、複数のプライマリノードを持つ Amazon EMR クラスターで動作するようになりました。

## 既知の問題

- ソートマージ結合の全体的なステージコード生成機能が改善されたことにより、有効にした場合、メモリプレッシャーが増加する可能性があります。この最適化によってパフォーマンスは向上しますが、十分なメモリを提供するように `spark.yarn.executor.memoryOverheadFactor` が調整されていない場合、ジョブの再試行または失敗が生じる可能性があります。この機能を無効にするには、`spark.sql.sortMergeJoinExec.extendedCodegen.enabled` を `false` に設定します。
- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.25.0

次のリリースノートには、Amazon EMR リリース 5.25.0 に関する情報が含まれています。5.24.1 からの変更が含まれています。

初回リリース日: 2019 年 7 月 17 日

最終更新日: 2019 年 10 月 30 日

Amazon EMR 5.25.0

アップグレード

- AWS SDK for Java 1.11.566

- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- Zookeeper 3.4.14

## 新機能

- (2019 年 10 月 30 日) Amazon EMR バージョン 5.25.0 以降では、コンソールでクラスターの [概要] ページまたは [アプリケーションの履歴] タブから Spark 履歴サーバー UI に接続できます。SSH 接続でウェブプロキシを設定する代わりに、Spark 履歴サーバー UI にすばやくアクセスしてアプリケーションのメトリクスを確認し、アクティブなクラスターと終了したクラスターに関連するログファイルにアクセスできます。詳細については、「Amazon EMR 管理ガイド」の「[永続アプリケーションユーザーインターフェイスへのクラスター外アクセス](#)」を参照してください。

## 変更、機能強化、解決した問題

- Spark
  - ブルームフィルターを使用して入力を事前フィルターすることで、一部の結合のパフォーマンスが改善されました。最適化はデフォルトでは無効になっており、Spark 設定パラメータ `spark.sql.bloomFilterJoin.enabled` を `true` に設定すると、有効にできます。
  - 文字列型の列によるグループ化のパフォーマンスが改善されました。
  - HBase がインストールされていないクラスターの R4 インスタンスタイプのデフォルトの Spark エグゼキューターメモリとコアの設定が改善されました。
  - プルーニングするテーブルが結合の左側にある必要がある、ダイナミックパーティションプルーニング機能に関する以前の問題が解決されました。
  - エイリアスを含む追加のケースに適用できるように、DISTINCT before INTERSECT 最適化が改善されました。
  - DISTINCT クエリが後に続く JOIN の SQL 計画統計推論が改善されました。この改善はデフォルトでは無効になっており、Spark 設定パラメータ `spark.sql.statsImprovements.enabled` を `true` に設定すると、有効にできます。この最適化は Distinct before Intersect 機能で必要であ

り、`spark.sql.optimizer.distinctBeforeIntersect.enabled` を `true` に設定すると、自動的に有効になります。

- テーブルのサイズとフィルターに基づいて結合順序が最適化されました。この最適化はデフォルトでは無効になっており、Spark 設定パラメータ `spark.sql.optimizer.sizeBasedJoinReorder.enabled` を `true` に設定すると、有効にできます。

詳細については、「[Spark のパフォーマンスの最適化](#)」を参照してください。

#### • EMRFS

- EMRFS の設定 `fs.s3.buckets.create.enabled` は、現在、デフォルトでは無効になります。テストでは、この設定を無効にすると、パフォーマンスが向上し、S3 バケットが意図せずに作成されることがなくなることがわかりました。アプリケーションがこの機能に依存する場合は、`emrfs-site` 設定分類でプロパティ `fs.s3.buckets.create.enabled` を `true` に設定すると、この機能を有効にできます。詳細については、「[クラスターの作成時に設定を指定する](#)」を参照してください。
- セキュリティ設定のローカルディスク暗号化と S3 暗号化の改善 (2019 年 8 月 5 日)
  - セキュリティ設定のセットアップで、Amazon S3 暗号化設定をローカルディスク暗号化設定から分離しました。
  - リリース 5.24.0 以降で、EBS 暗号化を有効にするオプションが追加されました。このオプションを選択すると、ストレージボリュームに加えてルートデバイスボリュームが暗号化されます。以前のバージョンでは、ルートデバイスボリュームを暗号化するにはカスタム AMI を使用する必要がありました。
  - 詳細については、「Amazon EMR 管理ガイド」の「[暗号化オプション](#)」を参照してください。

#### 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.24.1

次のリリースノートには、Amazon EMR リリース 5.24.1 に関する情報が含まれています。5.24.0 からの変更が含まれています。

初回リリース日: 2019 年 6 月 26 日

変更点、機能強化、解決した問題

- TCP SACK サービス拒否の問題 ([AWS-2019-005](#)) など、重要な Linux カーネルセキュリティ更新を含めるように、Amazon EMR のデフォルトの Amazon Linux AMI を更新しました。

既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは /etc/hadoop.keytab にあります。プリンシパルの形式は hadoop/<hostname>@<REALM> です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.24.0

次のリリースノートには、Amazon EMR リリース 5.24.0 に関する情報が含まれています。5.23.0 からの変更が含まれています。

初回リリース日: 2019 年 6 月 11 日

最終更新日: 2019 年 8 月 5 日

### アップグレード

- Flink 1.8.0
- Hue 4.4.0
- JupyterHub 0.9.6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK for Java 1.11.546



- コネクタおよびドライバー:
  - DynamoDB Connector 4.9.0
  - MariaDB Connector 2.4.1
  - Amazon Redshift JDBC ドライバー 1.2.27.1051

## 変更、機能強化、解決した問題

- Spark
  - パーティションを動的にプルーニングするための最適化を追加しました。この最適化はデフォルトで無効になっています。これを有効にするには、Spark 設定パラメータ `spark.sql.dynamicPartitionPruning.enabled` を `true` に設定します。
  - INTERSECT のクエリのパフォーマンスが改善されました。この最適化はデフォルトで無効になっています。これを有効にするには、Spark 設定パラメータ `spark.sql.optimizer.distinctBeforeIntersect.enabled` を `true` に設定します。
  - 同じリレーションを使用する集計でスカラーサブクエリをフラット化するための最適化が追加されました。この最適化はデフォルトで無効になっています。これを有効にするには、Spark 設定パラメータ `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` を `true` に設定します。
  - 全体的なステージコード生成が改善されました。

詳細については、「[Spark のパフォーマンスの最適化](#)」を参照してください。

- セキュリティ設定のローカルディスク暗号化と S3 暗号化の改善 (2019 年 8 月 5 日)
  - セキュリティ設定のセットアップで、Amazon S3 暗号化設定をローカルディスク暗号化設定から分離しました。
  - EBS 暗号化を有効にするオプションが追加されました。このオプションを選択すると、ストレージボリュームに加えてルートデバイスボリュームが暗号化されます。以前のバージョンでは、ルートデバイスボリュームを暗号化するにはカスタム AMI を使用する必要がありました。
  - 詳細については、「Amazon EMR 管理ガイド」の「[暗号化オプション](#)」を参照してください。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題



Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.23.0

次のリリースノートには、Amazon EMR リリース 5.23.0 に関する情報が含まれています。5.22.0 からの変更が含まれています。

初回リリース日: 2019 年 4 月 1 日

最終更新日: 2019 年 4 月 30 日

### アップグレード

- AWS SDK for Java 1.11.519

## 新機能

- (2019 年 4 月 30 日) Amazon EMR 5.23.0 以降では、3 つのプライマリノードを持つクラスターを起動して、YARN Resource Manager、HDFS、Spark NameNode、Hive、Ganglia などのアプリケーションの高可用性をサポートできます。プライマリノードは、現在この機能による潜在的な単一障害点ではありません。プライマリノードのいずれかに障害が発生した場合、Amazon EMR は自動的にスタンバイプライマリノードにフェイルオーバーし、障害が発生したプライマリノードを同じ設定とブートストラップアクションを持つ新しいプライマリノードに置き換えます。詳細については、「[プライマリノードの計画と設定](#)」を参照してください。

## 既知の問題

- Tez UI (Amazon EMR リリース 5.26.0 で修正)

Tez UI は、複数のプライマリノードを持つ EMR クラスターでは動作しません。

- Hue (Amazon EMR リリース 5.24.0 で修正)

- Amazon EMR で実行される Hue では、Solr はサポートされません。Amazon EMR リリース 5.20.0 以降、誤設定の問題により Solr が有効になり、次のような無害なエラーメッセージが表示されます。

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Solr のエラーメッセージが表示されないようにするには、以下の操作を行います。

1. SSH を使用してプライマリノードのコマンドラインに接続します。
2. テキストエディタを使用して、hue.ini ファイルを開きます。例:

```
sudo vim /etc/hue/conf/hue.ini
```

3. 語句 `appblacklist` を検索して、行を次のように変更します。

```
appblacklist = search
```

4. 変更を保存して、次の例に示すように Hue を再起動します。

```
sudo stop hue; sudo start hue
```

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.22.0

次のリリースノートには、Amazon EMR リリース 5.22.0 に関する情報が含まれています。5.21.0 からの変更が含まれています。

#### Important

Amazon EMR リリース 5.22.0 以降、Amazon EMR は AWS 署名バージョン 4 のみを使用して Amazon S3 へのリクエストを認証します。以前の Amazon EMR リリースでは、リリース

ノートで AWS Signature Version 4 のみが使用されていることが示されていない限り、場合によっては Signature Version 2 が使用されます。詳細については、Amazon Simple Storage Service デベロッパーガイドの「[リクエストの認証 \(AWS 署名バージョン 4\)](#)」および「[リクエストの認証 \(AWS 署名バージョン 2\)](#)」を参照してください。

初回リリース日: 2019 年 3 月 20 日

## アップグレード

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- コネクタおよびドライバー:
  - DynamoDB Connector 4.8.0
  - MariaDB Connector 2.2.6
  - Amazon Redshift JDBC ドライバー 1.2.20.1043

## 新機能

- EBS 専用ストレージを使用する EC2 インスタンスタイプのデフォルトの EBS 設定を変更しました。Amazon EMR リリース 5.22.0 以降を使用してクラスターを作成する場合、デフォルトの EBS ストレージ容量はインスタンスのサイズに基づいて増加します。さらに、増加したストレージを複数のボリュームに分割することで、IOPS のパフォーマンスが向上しました。別の EBS インスタンスストレージ設定を使用する場合は、EMR クラスターを作成する際、または既存のクラスターをノードに追加する際に指定することができます。インスタンスタイプごとにデフォルトで割り当てられるストレージ容量とボリューム数の詳細については、「Amazon EMR 管理ガイド」の「[インスタンスのデフォルト EBS ストレージ](#)」を参照してください。

## 変更、機能強化、解決した問題

- Spark
  - YARN の Spark 用に新しい設定プロパティ `spark.yarn.executor.memoryOverheadFactor` が導入されました。このプロパティの値

は、メモリアーヘッドの値をエグゼキュターメモリのパーセントに設定するスケール係数であり、最小は 384 MB です。spark.yarn.executor.memoryOverhead を使用してメモリアーヘッドが明示的に設定されている場合、このプロパティは影響しません。デフォルト値は 0.1875 で、18.75% を表します。Amazon EMR のこのデフォルトは、Spark で内部的に設定されたデフォルトの 10% よりも、エグゼキュターのメモリアーヘッド用に YARN コンテナに余分なスペースを残します。Amazon EMR のデフォルトの 18.75% では、TPC-DS ベンチマークでメモリ関連の障害が減ることが実証されています。

- パフォーマンスを向上させるために [SPARK-26316](#) をバックポートしました。
- Amazon EMR バージョン 5.19.0、5.20.0、および 5.21.0 では、YARN ノードラベルは HDFS ディレクトリに保存されます。状況によっては、このために、コアノードの起動が遅延し、クラスターのタイムアウトや起動エラーが発生します。Amazon EMR 5.22.0 以降、この問題は解決されています。YARN ノードラベルは各クラスターノードのローカルディスクに保存され、HDFS への依存が回避されています。

## 既知の問題

- Hue (Amazon EMR リリース 5.24.0 で修正)
  - Amazon EMR で実行される Hue では、Solr はサポートされません。Amazon EMR リリース 5.20.0 以降、誤設定の問題により Solr が有効になり、次のような無害なエラーメッセージが表示されます。

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Solr のエラーメッセージが表示されないようにするには、以下の操作を行います。

- SSH を使用してプライマリノードのコマンドラインに接続します。
- テキストエディタを使用して、hue.ini ファイルを開きます。例:

```
sudo vim /etc/hue/conf/hue.ini
```

- 語句 appblacklist を検索して、行を次のように変更します。

```
appblacklist = search
```

#### 4. 変更を保存して、次の例に示すように Hue を再起動します。

```
sudo stop hue; sudo start hue
```

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

##### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

##### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.21.1

次のリリースノートには、Amazon EMR リリース 5.21.1 に関する情報が含まれています。5.21.0 からの変更が含まれています。

初回リリース日: 2019 年 7 月 18 日

## 変更点、機能強化、解決した問題

- TCP SACK サービス拒否の問題 ([AWS-2019-005](#)) など、重要な Linux カーネルセキュリティ更新を含めるように、Amazon EMR のデフォルトの Amazon Linux AMI を更新しました。

## 既知の問題

- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

### 回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに hadoop ユーザーとして SSH 接続します。
- 次のコマンドを実行して hadoop ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.21.0

次のリリースノートには、Amazon EMR リリース 5.21.0 に関する情報が含まれています。5.20.0 からの変更が含まれています。

リリース日: 2019 年 2 月 18 日

最終更新日: 2019 年 4 月 3 日

## アップグレード

- Flink 1.7.0
- Presto 0.215
- AWS SDK for Java 1.11.479

## 新機能

- (2019 年 4 月 3 日) Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## 変更、機能強化、解決した問題

- Zeppelin
  - [ZEPPELIN-3878](#) をバックポートしました。

## 既知の問題

- Hue (Amazon EMR リリース 5.24.0 で修正)
  - Amazon EMR で実行される Hue では、Solr はサポートされません。Amazon EMR リリース 5.20.0 以降、誤設定の問題により Solr が有効になり、次のような無害なエラーメッセージが表示されます。

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Solr のエラーメッセージが表示されないようにするには、以下の操作を行います。



1. SSH を使用してプライマリノードのコマンドラインに接続します。
2. テキストエディタを使用して、hue.ini ファイルを開きます。例:

```
sudo vim /etc/hue/conf/hue.ini
```

3. 語句 `appblacklist` を検索して、行を次のように変更します。

```
appblacklist = search
```

4. 変更を保存して、次の例に示すように Hue を再起動します。

```
sudo stop hue; sudo start hue
```

- Tez

- この問題は、Amazon EMR 5.22.0 で修正されています。

クラスタープライマリノードへの SSH 接続を使用して `http://MasterDNS:8080/tez-ui` で Tez UI に接続すると、「Adapter operation failed - Timeline server (ATS) is out of reach. Either it is down, or CORS is not enabled」(アダプタの操作に失敗しました - タイムラインサーバー (ATS) に到達できません。停止しているか、または CORS が有効になっていません) というエラーが表示されるか、タスクで予期せずに N/A と表示されます。

これは、Tez UI がプライマリノードのホスト名ではなく `localhost` を使用して YARN タイムラインサーバーにリクエストを行ったことが原因で発生します。回避策として、ブートストラップアクションまたはステップとしてスクリプトを実行できます。このスクリプトで、`Tez configs.env` ファイル内のホスト名を更新します。スクリプトの詳細と場所については、「[Bootstrap Instructions](#)」を参照してください。

- Amazon EMR バージョン 5.19.0、5.20.0、および 5.21.0 では、YARN ノードラベルは HDFS ディレクトリに保存されます。状況によっては、このために、コアノードの起動が遅延し、クラスターのタイムアウトや起動エラーが発生します。Amazon EMR 5.22.0 以降、この問題は解決されています。YARN ノードラベルは各クラスターノードのローカルディスクに保存され、HDFS への依存が回避されています。
- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエスト

と送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.20.0

次のリリースノートには、Amazon EMR リリース 5.20.0 に関する情報が含まれています。5.19.0 からの変更が含まれています。

初回リリース日: 2018 年 12 月 18 日

最終更新: 2019 年 1 月 22 日

### アップグレード

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214

- Spark 2.4.0
- TensorFlow 1.12.0
- Tez 0.9.1
- AWS SDK for Java 1.11.461

## 新機能

- (2019 年 1 月 22 日) Amazon EMR の Kerberos は、外部 KDC からのプリンシパルの認証をサポートするように改善されました。これにより、複数のクラスターが単一の外部 KDC を共有できるため、プリンシパル管理が集中化されます。さらに、外部 KDC は Active Directory ドメインとのクロス領域信頼を得られます。これにより、すべてのクラスターが Active Directory からプリンシパルを認証できます。詳しくは、「Amazon EMR 管理ガイド」の「[Kerberos 認証を使用する](#)」を参照してください。

## 変更、機能強化、解決した問題

- Amazon EMR のデフォルトの Amazon Linux AMI
  - Python3 パッケージが python 3.4 から 3.6 にアップグレードされました。
- EMRFS S3 向けに最適化されたコミッター
  - 現在、EMRFS S3 向けに最適化されたコミッターがデフォルトで有効になるため、書き込みパフォーマンスが向上します。詳細については、「[EMRFS S3 向けに最適化されたコミッターを使用する](#)」を参照してください。
- [Hive]
  - [HIVE-16686](#) をバックポートしました。
- Spark と Hive での Glue
  - EMR 5.20.0 以降では、AWS Glue Data Catalog をメタストアとして使用すると、Spark と Hive の並列パーティションプルーニングが自動的に有効になります。この変更により、複数のリクエストを並列に実行してパーティションを取得できるため、クエリ計画時間が大幅に短縮されます。同時に実行できるセグメントの総数は、1~10 の範囲です。デフォルト値は 5 であり、これが推奨される設定です。hive-site 設定分類の aws.glue.partition.num.segments プロパティを指定すると、これを変更できます。スロットリングが発生した場合は、値を 1 に変更して機能をオフにすることができます。詳細については、「[AWS Glue セグメント構造](#)」を参照してください。

## 既知の問題

- Hue (Amazon EMR リリース 5.24.0 で修正)
  - Amazon EMR で実行される Hue では、Solr はサポートされません。Amazon EMR リリース 5.20.0 以降、誤設定の問題により Solr が有効になり、次のような無害なエラーメッセージが表示されます。

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Solr のエラーメッセージが表示されないようにするには、以下の操作を行います。

1. SSH を使用してプライマリノードのコマンドラインに接続します。
2. テキストエディタを使用して、hue.ini ファイルを開きます。例:

```
sudo vim /etc/hue/conf/hue.ini
```

3. 語句 `appblacklist` を検索して、行を次のように変更します。

```
appblacklist = search
```

4. 変更を保存して、次の例に示すように Hue を再起動します。

```
sudo stop hue; sudo start hue
```

- Tez
  - この問題は、Amazon EMR 5.22.0 で修正されています。

クラスタープライマリノードへの SSH 接続を使用して `http://MasterDNS:8080/tez-ui` で Tez UI に接続すると、「Adapter operation failed - Timeline server (ATS) is out of reach. Either it is down, or CORS is not enabled」(アダプタの操作に失敗しました - タイムラインサーバー (ATS) に到達できません。停止しているか、または CORS が有効になっていません) というエラーが表示されるか、タスクで予期せずに N/A と表示されます。

これは、Tez UI がプライマリノードのホスト名ではなく `localhost` を使用して YARN タイムラインサーバーにリクエストを行ったことが原因で発生します。回避策として、ブートスト

ラップアクションまたはステップとしてスクリプトを実行できます。このスクリプトで、`Tez configs.env` ファイル内のホスト名を更新します。スクリプトの詳細と場所については、「[Bootstrap Instructions](#)」を参照してください。

- Amazon EMR バージョン 5.19.0、5.20.0、および 5.21.0 では、YARN ノードラベルは HDFS ディレクトリに保存されます。状況によっては、このために、コアノードの起動が遅延し、クラスターのタイムアウトや起動エラーが発生します。Amazon EMR 5.22.0 以降、この問題は解決されています。YARN ノードラベルは各クラスターノードのローカルディスクに保存され、HDFS への依存が回避されています。
- 複数のプライマリノードと Kerberos 認証を使用するクラスターの既知の問題

Amazon EMR リリース 5.20.0 以降で複数のプライマリノードと Kerberos 認証を使用してクラスターを実行すると、クラスターをしばらく実行した後で、スケールダウンやステップの送信などのクラスターオペレーションに問題が発生する可能性があります。期間は、定義した Kerberos チケットの有効期間によって異なります。スケールダウンの問題は、自動スケールダウンリクエストと送信した明示的なスケールダウンリクエストの両方に影響します。その他のクラスターオペレーションも影響を受ける可能性があります。

回避方法:

- 複数のプライマリノードを持つ EMR クラスターのリードプライマリノードに `hadoop` ユーザーとして SSH 接続します。
- 次のコマンドを実行して `hadoop` ユーザーの Kerberos チケットを更新します。

```
kinit -kt <keytab_file> <principal>
```

通常、キータブファイルは `/etc/hadoop.keytab` にあります。プリンシパルの形式は `hadoop/<hostname>@<REALM>` です。

#### Note

この回避策は、Kerberos チケットが有効になっている期間、効果があります。この期間はデフォルトで 10 時間ですが、Kerberos の設定で構成できます。Kerberos チケットの有効期限が切れたら、上記のコマンドを再実行する必要があります。

## リリース 5.19.0

次のリリースノートには、Amazon EMR リリース 5.19.0 に関する情報が含まれています。5.18.0 からの変更が含まれています。

リリース日: 2018 年 11 月 7 日

最終更新日: 2018 年 11 月 19 日

### アップグレード

- Hadoop 2.8.5
- Flink 1.6.1
- JupyterHub 0.9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1.11.0
- Zookeeper 3.4.13
- AWS SDK for Java 1.11.433

### 新機能

- (2018 年 11 月 19 日) EMR Notebooks は、Jupyter Notebook に基づいたマネージド環境です。、Spark SQL、Spark R PySpark、Scala の Spark マジックカーネルをサポートします。EMR Notebooks は、Amazon EMR リリース 5.18.0 以降を使用して作成されたクラスターで使用できます。詳細については、「Amazon EMR 管理ガイド」の「[EMR Notebooks の使用](#)」を参照してください。
- EMRFS S3 向けに最適化されたコミッターは、Spark と EMRFS を使用して Parquet ファイルを書き込むときに利用できます。このコミッターにより、書き込みパフォーマンスが向上します。詳細については、「[EMRFS S3 向けに最適化されたコミッターを使用する](#)」を参照してください。

### 変更、機能強化、解決した問題

- YARN
  - アプリケーションマスタープロセスをコアノードで実行するように制限するロジックを修正しました。現在、この機能では、yarn-site 設定分類および capacity-scheduler 設定分類の YARN ノードラベル機能とプロパティを使用しています。詳細については、「<https://>

[docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN.](https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN.)」を参照してください。

- Amazon EMR のデフォルトの Amazon Linux AMI
  - ruby18、php56、および gcc48 は、デフォルトではインストールされなくなりました。必要に応じて、yum を使用してインストールできます。
  - aws-sdk ruby gem は、デフォルトではインストールされなくなりました。必要に応じて、gem install aws-sdk を使用してインストールできます。特定のコンポーネントをインストールすることもできます。例えば gem install aws-sdk-s3 です。

## 既知の問題

- EMR Notebooks — 状況によっては、複数のノートブックエディタが開いているときに、ノートブックエディタがクラスターに接続できないように見えることがあります。このような場合は、ブラウザの Cookie をクリアしてから、ノートブックエディタを再度開きます。
- CloudWatch ContainerPending メトリクスと自動スケーリング — (5.20.0 で修正) Amazon EMR は に負の値を出力することがあります ContainerPending。ContainerPending が自動スケーリングルールで使用されている場合、自動スケーリングは期待どおりに動作しません。ContainerPending を自動スケーリングで使用しないでください。
- Amazon EMR バージョン 5.19.0、5.20.0、および 5.21.0 では、YARN ノードラベルは HDFS ディレクトリに保存されます。状況によっては、このために、コアノードの起動が遅延し、クラスターのタイムアウトや起動エラーが発生します。Amazon EMR 5.22.0 以降、この問題は解決されています。YARN ノードラベルは各クラスターノードのローカルディスクに保存され、HDFS への依存が回避されています。

## リリース 5.18.0

次のリリースノートには、Amazon EMR リリース 5.18.0 に関する情報が含まれています。5.17.0 からの変更が含まれています。

初回リリース日: 2018 年 10 月 24 日

### アップグレード

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210

- Spark 2.3.2
- Zeppelin 0.8.0

## 新機能

- Amazon EMR 5.18.0 以降では、Amazon EMR アーティファクトリポジトリを使用し、特定の Amazon EMR リリースで利用可能なライブラリおよび依存関係の正確なバージョンに対してジョブコードをビルドできます。詳細については、「[Amazon EMR アーティファクトリポジトリを使用して依存関係を確認する](#)」を参照してください。

## 変更、機能強化、解決した問題

- [Hive]
  - S3 Select のサポートが追加されました。詳細については、「[S3 Select と Hive を使用してパフォーマンスを向上させる](#)」を参照してください。
- Presto
  - [S3 Select Pushdown](#) のサポートが追加されました。詳細については、「[S3 Select Pushdown と Presto を使用してパフォーマンスを向上させる](#)」を参照してください。
- Spark
  - Spark のデフォルトの log4j 設定が、Spark ストリーミングジョブのコンテナログを毎時ロールするように変更されました。これにより、長時間実行されている Spark ストリーミングジョブのログが削除されるのを防ぐことができます。

## リリース 5.17.1

次のリリースノートには、Amazon EMR リリース 5.17.1 に関する情報が含まれています。5.17.0 からの変更が含まれています。

初回リリース日: 2019 年 7 月 18 日

## 変更点、機能強化、解決した問題

- TCP SACK サービス拒否の問題 ([AWS-2019-005](#)) など、重要な Linux カーネルセキュリティ更新を含めるように、Amazon EMR のデフォルトの Amazon Linux AMI を更新しました。



## リリース 5.17.0

次のリリースノートには、Amazon EMR リリース 5.17.0 に関する情報が含まれています。5.16.0 からの変更が含まれています。

初回リリース日: 2018 年 8 月 30 日

### アップグレード

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

### 新機能

- TensorFlow のサポートが追加されました。詳細については、「[TensorFlow](#)」を参照してください。

### 変更点、機能強化、解決した問題

- JupyterHub
  - Amazon S3 でのノートブックの永続性のサポートが追加されました。詳細については、「[Amazon S3 でノートブックの永続性を設定するには](#)」を参照してください。
- Spark
  - [S3 Select](#) のサポートが追加されました。詳細については、「[S3 Select と Spark を使用してクエリパフォーマンスを向上させる](#)」を参照してください。
- Amazon EMR バージョン 5.14.0、5.15.0、または 5.16.0 での Cloudwatch メトリクスと自動スケーリング機能に関する問題が解決されました。

### 既知の問題

- Livy がインストールされた Kerberized クラスターを作成すると、Livy が失敗して、簡易認証が有効になっていないというエラーが表示されます。Livy サーバーを再起動すると、問題が解決されます。回避策として、クラスターの作成時に、プライマリノードで `sudo restart livy-server` を実行するステップを追加します。

- 作成日が 2018-08-11 の Amazon Linux AMI に基づくカスタム Amazon Linux AMI を使用すると、Oozie サーバーの起動に失敗します。Oozie を使用する場合は、作成日が異なる Amazon Linux AMI ID に基づいてカスタム AMI を作成します。次の AWS CLI コマンドを使用して、2018.03 バージョンのすべての HVM Amazon Linux AMIs のイメージ IDs のリストをリリース日とともに返すことができます。これにより、ベースとして適切な Amazon Linux AMI を選択できます。を us-west-2 などのリージョン識別子 MyRegion に置き換えます。

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?Name!=`null`][?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].[CreationDate,ImageId,Name]' --output text | sort -rk1
```

## リリース 5.16.0

次のリリースノートには、Amazon EMR リリース 5.16.0 に関する情報が含まれています。5.15.0 からの変更が含まれています。

初回リリース日: 2018 年 7 月 19 日

### アップグレード

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK for Java 1.11.336
- CUDA 9.2
- Redshift JDBC ドライバー 1.2.15.1025

### 変更、機能強化、解決した問題

- HBase
  - [HBASE-20723](#) をバックポートしました。

- Presto
  - LDAP 認証をサポートするための設定変更。詳細については、「[Amazon EMR の Presto に LDAP 認証を使用する](#)」を参照してください。
- Spark
  - Apache Spark バージョン 2.3.1 は Amazon EMR リリース 5.16.0 以降で利用でき、[CVE-2018-8024](#) および [CVE-2018-1334](#) に対処します。Spark の以前のバージョンを Spark バージョン 2.3.1 以降に移行することをお勧めします。

## 既知の問題

- このリリースバージョンでは、c1.medium インスタンスタイプまたは m1.small インスタンスタイプはサポートされていません。これらのインスタンスタイプのいずれかを使用するクラスターは起動しません。回避策として、別のインスタンスタイプを指定するか、別のリリースバージョンを使用します。
- Livy がインストールされた Kerberized クラスターを作成すると、Livy が失敗して、簡易認証が有効になっていないというエラーが表示されます。Livy サーバーを再起動すると、問題が解決されます。回避策として、クラスターの作成時に、プライマリノードで `sudo restart livy-server` を実行するステップを追加します。
- プライマリノードの再起動またはインスタンスコントローラーの再起動後、CloudWatch メトリクスは収集されず、自動スケーリング機能は Amazon EMR バージョン 5.14.0、5.15.0、または 5.16.0 では使用できません。この問題は、Amazon EMR 5.17.0 で修正されています。

## リリース 5.15.0

次のリリースノートには、Amazon EMR リリース 5.15.0 に関する情報が含まれています。5.14.0 からの変更が含まれています。

初回リリース日: 2018 年 6 月 21 日

### アップグレード

- HBase を 1.4.4 にアップグレードしました
- Hive を 2.3.3 にアップグレードしました
- Hue を 4.2.0 にアップグレードしました
- Oozie を 5.0.0 にアップグレードしました

- ZooKeeper を 3.4.12 にアップグレードしました
- AWS SDK を 1.11.333 にアップグレードしました

#### 変更点、機能強化、解決した問題

- [Hive]
  - [HIVE-18069](#) をバックポートしました。
- Hue
  - Kerberos が有効になっている場合に Livy で正しく認証されるように Hue を更新しました。Amazon EMR で Kerberos を使用する場合、Livy がサポートされるようになりました。
- JupyterHub
  - Amazon EMR がデフォルトで LDAP クライアントライブラリをインストールする JupyterHub ように更新されました。
  - 自己署名証明書を生成するスクリプトのエラーを修正しました。

#### 既知の問題

- このリリースバージョンでは、c1.medium インスタンスタイプまたは m1.small インスタンスタイプはサポートされていません。これらのインスタンスタイプのいずれかを使用するクラスターは起動しません。回避策として、別のインスタンスタイプを指定するか、別のリリースバージョンを使用します。
- プライマリノードの再起動またはインスタンスコントローラーの再起動後、CloudWatch メトリクスは収集されず、自動スケーリング機能は Amazon EMR バージョン 5.14.0、5.15.0、または 5.16.0 では使用できません。この問題は、Amazon EMR 5.17.0 で修正されています。

## リリース 5.14.1

次のリリースノートには、Amazon EMR リリース 5.14.1 に関する情報が含まれています。5.14.0 からの変更が含まれています。

初回リリース日: 2018 年 10 月 17 日

潜在的なセキュリティ脆弱性に対処するために Amazon EMR のデフォルトの AMI を更新しました。

## リリース 5.14.0

次のリリースノートには、Amazon EMR リリース 5.14.0 に関する情報が含まれています。5.13.0 からの変更が含まれています。

初回リリース日: 2018 年 6 月 4 日

### アップグレード

- Apache Flink を 1.4.2 にアップグレードしました
- Apache MXNet を 1.1.0 にアップグレードしました
- Apache Sqoop を 1.4.7 にアップグレードしました

### 新機能

- JupyterHub サポートを追加しました。詳細については、「[JupyterHub](#)」を参照してください。

### 変更、機能強化、解決した問題

- EMRFS
  - Amazon S3 へのリクエスト内の UserAgent 文字列が、呼び出し元のプリンシパルのユーザーおよびグループの情報を含めるように更新されました。これは、より包括的なリクエスト追跡のための AWS CloudTrail ログで使用できます。
- HBase
  - [HBASE-20447](#) が含まれました。これにより、特に分割されたリージョンで、キャッシュの問題を引き起こす可能性のある問題に対処します。
- MXNet
  - OpenCV ライブラリを追加しました。
- Spark
  - Spark が EMRFS を使用して Amazon S3 の場所に Parquet ファイルを書き込むと、FileOutputCommitter アルゴリズムがバージョン 1 ではなくバージョン 2 を使用するように更新されました。これにより、名前変更の数が減り、アプリケーションのパフォーマンスが向上します。この変更は、以下には影響しません。
    - Spark 以外のアプリケーション。
    - HDFS など、他のファイルシステムに書き込むアプリケーション ( のバージョン 1 は引き続き使用 FileOutputCommitter ) 。

- テキストや csv など、EMRFS 直接書き込みを既に使用している、他の出力形式を使用するアプリケーション。

## 既知の問題

### • JupyterHub

- クラスターの作成時に設定分類を使用して Jupyter Notebook JupyterHub と個々の Jupyter Notebook をセットアップすることはサポートされていません。各ユーザーの `jupyterhub_config.py` ファイルと `jupyter_notebook_config.py` ファイルを手動で編集します。詳細については、「[の設定 JupyterHub](#)」を参照してください。
- JupyterHub は、プライベートサブネット内のクラスターで起動に失敗し、メッセージで失敗します `Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'`。これは、自己署名証明書を生成するスクリプトのエラーが原因で発生します。次の回避策を使用して自己署名証明書を生成します。すべてのコマンドは、プライマリノードに接続している間に実行されます。

1. 証明書生成スクリプトをコンテナからプライマリノードにコピーします。

```
sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
```

2. テキストエディタを使用して 23 行目を変更し、以下に示すようにパブリックホスト名をローカルホスト名に変更します。

```
local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
```

3. スクリプトを実行して、自己署名証明書を生成します。

```
sudo bash ./gen_self_signed_cert.sh
```

4. スクリプトによって生成される証明書ファイルを `/etc/jupyter/conf/` ディレクトリに移動します。:

```
sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
```

`tail jupyter.log` ファイルを使用して、`g` JupyterHub 再起動し、200 レスポンスコードを返していることを確認できます。例:

```
tail -f /var/log/jupyter/jupyter.log
```

これで、次のようなレスポンスが返されます。

```
# [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
https://:9443/
# 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
```

- プライマリノードの再起動またはインスタンスコントローラーの再起動後、CloudWatch メトリクスは収集されず、自動スケーリング機能は Amazon EMR バージョン 5.14.0、5.15.0、または 5.16.0 では使用できません。この問題は、Amazon EMR 5.17.0 で修正されています。

## リリース 5.13.0

次のリリースノートには、Amazon EMR リリース 5.13.0 に関する情報が含まれています。5.12.0 からの変更が含まれています。

### アップグレード

- Spark を 2.3.0 にアップグレードしました
- HBase を 1.4.2 にアップグレードしました
- Presto を 0.194 にアップグレードしました
- AWS SDK for Java を 1.11.297 にアップグレードしました

### 変更点、機能強化、解決した問題

- [Hive]
  - [HIVE-15436](#) をバックポートしました。ビューのみを返すように Hive API が拡張されました。

### 既知の問題

- MXNet には現在 OpenCV ライブラリがありません。

## リリース 5.12.2

次のリリースノートには、Amazon EMR リリース 5.12.2 に関する情報が含まれています。5.12.1 からの変更が含まれています。

初回リリース日: 2018 年 8 月 29 日

## 変更、機能強化、解決した問題

- このリリースでは、潜在的なセキュリティ脆弱性に対処しています。

## リリース 5.12.1

次のリリースノートには、Amazon EMR リリース 5.12.1 に関する情報が含まれています。5.12.0 からの変更が含まれています。

初回リリース日: 2018 年 3 月 29 日

## 変更、機能強化、解決した問題

- 潜在的な脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。

## リリース 5.12.0

次のリリースノートには、Amazon EMR リリース 5.12.0 に関する情報が含まれています。5.11.1 からの変更が含まれています。

## アップグレード

- AWS SDK for Java 1.11.238 Pub 1.11.267。詳細については、の [AWS SDK for Java 変更ログ](#) を参照してください GitHub。
- Hadoop 2.7.3 ⇒ 2.8.3。詳細については、「[Apache Hadoop リリース](#)」を参照してください。
- Flink 1.3.2 ⇒ 1.4.0。詳細については、「[Apache Flink 1.4.0 リリースに関するお知らせ](#)」を参照してください。
- HBase 1.3.1 ⇒ 1.4.0。詳細については、「[HBase リリースに関するお知らせ](#)」を参照してください。
- Hue 4.0.1 ⇒ 4.1.0。詳細については、「[リリースノート](#)」を参照してください。
- MxNet 0.12.0 Gbps 1.0.0。詳細については、「[」のMXNet 変更ログ](#)」を参照してください GitHub。
- Presto 0.187 ⇒ 0.188。詳細については、「[リリースノート](#)」を参照してください。



## 変更、機能強化、解決した問題

- Hadoop
  - `yarn.resourcemanager.decommissioning.timeout` プロパティは `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs` に変更されました。このプロパティを使用して、クラスターのスケールダウンをカスタマイズできます。詳細については、「Amazon EMR 管理ガイド」の「[クラスターのスケールダウン](#)」を参照してください。
  - Hadoop CLI で、直接コピーを指定する `-d` オプションが `cp (copy)` コマンドに追加されました。これを使用すると、中間の `.COPYING` ファイルを作成しないようにできるため、Amazon S3 間でのデータのコピーが高速になります。詳細については、[HADOOP-12384](#) を参照してください。
- Pig
  - Pig 環境プロパティの設定を簡素化する `pig-env` 構成分類が追加されました。詳細については、「[アプリケーションの設定](#)」を参照してください。
- Presto
  - Presto の `presto-connector-redshift` 設定ファイルで値を指定するために使用できる、`redshift.properties` 設定分類が追加されました。詳細については、Presto のドキュメントの「[Redshift Connector](#)」、および「[アプリケーションの設定](#)」を参照してください。
  - EMRFS の Presto サポートが追加されました。これがデフォルト設定です。以前の Amazon EMR リリースでは、唯一のオプションである `PrestoS3FileSystem` が使用されていました。詳細については、「[EMRFS と PrestoS3FileSystem の設定](#)」を参照してください。

### Note

Amazon EMR バージョン 5.12.0 を使用して Amazon S3 で基盤となるデータをクエリするとき、Presto エラーが発生することがあります。これは、Presto が `emrfs-site.xml` から設定分類値を取得できないためです。回避策として、`usr/lib/presto/plugin/hive-hadoop2/` の下に `emrfs` サブディレクトリを作成し、既存の `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` ファイルへのシンボリックリンクを `usr/lib/presto/plugin/hive-hadoop2/emrfs` に作成します。次に `presto-server` プロセスを再起動します (`sudo presto-server stop` に続いて `sudo presto-server start`)。

- Spark
  - バックポートされた [SPARK-22036: BigDecimal multiplication は、null を返すことがあります](#)。

## 既知の問題

- MXNet には OpenCV ライブラリは含まれていません。
- デフォルトでは R はクラスターノードにインストールされないため、カスタム AMI を使用して作成されたクラスターで SparkR を使用することはできません。

## リリース 5.11.3

次のリリースノートには、Amazon EMR リリース 5.11.3 に関する情報が含まれています。5.11.2 からの変更が含まれています。

初回リリース日: 2019 年 7 月 18 日

### 変更点、機能強化、解決した問題

- TCP SACK サービス拒否の問題 ([AWS-2019-005](#)) など、重要な Linux カーネルセキュリティ更新を含めるように、Amazon EMR のデフォルトの Amazon Linux AMI を更新しました。

## リリース 5.11.2

次のリリースノートには、Amazon EMR リリース 5.11.2 に関する情報が含まれています。5.11.1 からの変更が含まれています。

初回リリース日: 2018 年 8 月 29 日

### 変更、機能強化、解決した問題

- このリリースでは、潜在的なセキュリティ脆弱性に対処しています。

## リリース 5.11.1

次のリリースノートには、Amazon EMR バージョン 5.11.1 リリースに関する情報が含まれています。Amazon EMR 5.11.0 リリースからの変更が含まれています。

初回リリース日: 2018 年 1 月 22 日

### 変更、機能強化、解決した問題

- 投機的実行 (CVE-2017-5715、CVE-2017-5753、CVE-2017-5754) に関連する脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しまし

た。詳細については、「<https://aws.amazon.com/security/security-bulletins/AWS-2018-013/>」を参照してください。

## 既知の問題

- MXNet には OpenCV ライブラリは含まれていません。
- Hive 2.3.2 では `hive.compute.query.using.stats=true` がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、`hive.compute.query.using.stats=true` が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して `SELECT COUNT(*)` クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、`ANALYZE TABLE` コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

## リリース 5.11.0

次のリリースノートには、Amazon EMR バージョン 5.11.0 リリースに関する情報が含まれています。Amazon EMR 5.10.0 リリースからの変更が含まれています。

### アップグレード

このリリースでは、以下のアプリケーションおよびコンポーネントがアップグレードされ、以下のバージョンが含まれています。

- Hive 2.3.2
- Spark 2.2.1
- SDK for Java 1.11.238

### 新機能

- Spark
  - `spark.decommissioning.timeout.threshold` 設定が追加されました。スポットインスタンス使用時の Spark 廃棄の動作が向上します。詳細については、「[ノード停止の動作設定](#)」を参照してください。

- `aws-sagemaker-spark-sdk` コンポーネントを Spark に追加しました。これにより、Amazon SageMaker Spark と Spark の統合に関連する依存関係がインストールされます [SageMaker](#)。Amazon SageMaker Spark を使用して、Amazon SageMaker ステージを使用して Spark 機械学習 (ML) パイプラインを構築できます。詳細については、「Amazon デベロッパーガイド」の「での [SageMaker Spark readme GitHub](#)」と「[Amazon での Apache Spark の使用 SageMaker](#)」を参照してください。 SageMaker

## 既知の問題

- MXNet には OpenCV ライブラリは含まれていません。
- Hive 2.3.2 では `hive.compute.query.using.stats=true` がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、`hive.compute.query.using.stats=true` が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して `SELECT COUNT(*)` クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、`ANALYZE TABLE` コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

## リリース 5.10.0

次のリリースノートには、Amazon EMR バージョン 5.10.0 リリースに関する情報が含まれています。Amazon EMR 5.9.0 リリースからの変更が含まれています。

### アップグレード

このリリースでは、以下のアプリケーションおよびコンポーネントがアップグレードされ、以下のバージョンが含まれています。

- AWS SDK for Java 1.11.221
- Hive 2.3.1
- Presto 0.187

## 新機能

- Kerberos 認証のサポートが追加されました。詳しくは、「Amazon EMR 管理ガイド」の「[Kerberos 認証を使用する](#)」を参照してください。
- Amazon S3 への EMRFS リクエストの IAM ロールのサポートを追加しました。詳細については、「Amazon EMR 管理ガイド」の「[Amazon S3 への EMRFS リクエストの IAM ロールを設定する](#)」を参照してください。
- GPU ベースの P2 および P3 のインスタンスタイプのサポートが追加されました。詳細については、「[Amazon EC2 P2 インスタンス](#)」および「[Amazon EC2 P3 インスタンス](#)」を参照してください。これらのインスタンスタイプには、デフォルトで NVIDIA ドライバー 384.81 および CUDA ドライバー 9.0.176 がインストールされています。
- [Apache MXNet](#) のサポートが追加されました。

## 変更、機能強化、解決した問題

- Presto
  - AWS Glue データカタログをデフォルトの Hive メタストアとして使用するためのサポートが追加されました。詳細については、「[AWS Glue データカタログでの Presto の使用](#)」を参照してください。
  - 「[地理空間機能](#)」のサポートが追加されました。
  - 結合で「[ディスクに書き出す](#)」ことができるようになりました。
  - 「[Redshift コネクタ](#)」のサポートが追加されました。
- Spark
  - 「[SPARK-20640](#)」をバックポートしました。これにより、spark.shuffle.registration.timeout プロパティおよび spark.shuffle.registration.maxAttempts プロパティを使用して、rpc タイムアウトや、シャッフル登録値の再試行回数が設定可能になりました。
  - [SPARK-21549](#) をバックポートしました。これにより、HDFS 以外の場所にカスタム OutputFormat を書き込むときに発生するエラーが修正されます。
  - 「[Hadoop-13270](#)」をバックポートしました。
- Numpy、Scipy、Matplotlib の各ライブラリは、Amazon EMR の基本 AMI から削除されています。アプリケーションでこれらのライブラリが必要な場合、アプリケーションリポジトリで使用できるため、ブートストラップアクションで yum install を使用してすべてのノードにインストールすることができます。

- Amazon EMR の基本 AMI にアプリケーションの RPM パッケージが含まれなくなったため、その RPM パッケージはクラスターノードに存在しなくなりました。カスタム AMI と Amazon EMR の基本 AMI で、Amazon S3 の RPM パッケージリポジトリを参照できるようになりました。
- Amazon EC2 で 1 秒単位の請求が導入されたため、デフォルトの [スケールダウン動作] は [インスタンス時間で削除する] ではなく [タスク完了で削除する] になっています。詳細については、「[クラスターのスケールダウンを設定する](#)」を参照してください。

## 既知の問題

- MXNet には OpenCV ライブラリは含まれていません。
- Hive 2.3.1 では `hive.compute.query.using.stats=true` がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、`hive.compute.query.using.stats=true` が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して `SELECT COUNT(*)` クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、`ANALYZE TABLE` コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

## リリース 5.9.0

次のリリースノートには、Amazon EMR バージョン 5.9.0 リリースに関する情報が含まれています。Amazon EMR 5.8.0 リリースからの変更が含まれています。

リリース日: 2017 年 10 月 5 日

最新機能更新日: 2017 年 10 月 12 日

## アップグレード

このリリースでは、以下のアプリケーションおよびコンポーネントがアップグレードされ、以下のバージョンが含まれています。

- AWS SDK for Java バージョン 1.11.183
- Flink 1.3.2

- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

## 新機能

- Livy サポート (バージョン 0.4.0-incubating) を追加しました。詳細については、「[Apache Livy](#)」を参照してください。
- Spark の Hue ノートブックのサポートを追加しました。
- i3-シリーズ Amazon EC2 インスタンスのサポートを追加しました (2017 年 10 月 12 日)。

## 変更、機能強化、解決した問題

- Spark
  - 手動のサイズ変更または自動のスケーリングポリシーのリクエストによるノードの終了処理を、Spark がより適切に行う、新しい機能のセットを追加しました。詳細については、「[ノード停止の動作設定](#)」を参照してください。
  - ブロック転送サービスの転送時の暗号化には 3DES に代わり SSL を使用します。これにより AES-NI で Amazon EC2 インスタンスタイプを使用する場合にパフォーマンスが向上します。
  - [SPARK-21494](#) を移植しました。
- Zeppelin
  - [ZEPPELIN-2377](#) を移植しました。
- HBase
  - パッチ [HBASE-18533](#) を追加しました。これにより、設定分類を使用して HBase BucketCache hbase-site設定に追加の値を指定できます。
- Hue
  - Hue の Hive クエリエディタに対する AWS Glue Data Catalog サポートが追加されました。
  - デフォルトでは、Hue のスーパーユーザーは、Amazon EMR の IAM ロールがアクセスを許可されているすべてのファイルにアクセスできます。新しく作成されたユーザーには、Amazon S3 ファイルブラウザへのアクセス許可は自動的に付与されません。グループに対して `filebrowser.s3_access` アクセス許可を有効にする必要があります。
- AWS Glue Data Catalog を使って作成された、基盤となる JSON データがアクセスできなくなる問題を解決しました。

## 既知の問題

- すべてのアプリケーションがインストールされ、デフォルトの Amazon EBS ルートボリュームサイズが変更されていない場合、クラスターの起動は失敗します。回避策として、の `aws emr create-cluster` コマンドを使用し AWS CLI、より大きな `--ebs-root-volume-size` パラメータを指定します。
- Hive 2.3.0 では `hive.compute.query.using.stats=true` がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、`hive.compute.query.using.stats=true` が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して `SELECT COUNT(*)` クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、`ANALYZE TABLE` コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

## リリース 5.8.2

次のリリースノートには、Amazon EMR リリース 5.8.2 に関する情報が含まれています。5.8.1 からの変更が含まれています。

初回リリース日: 2018 年 3 月 29 日

変更、機能強化、解決した問題

- 潜在的な脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。

## リリース 5.8.1

次のリリースノートには、Amazon EMR バージョン 5.8.1 リリースに関する情報が含まれています。Amazon EMR 5.8.0 リリースからの変更が含まれています。

初回リリース日: 2018 年 1 月 22 日



## 変更、機能強化、解決した問題

- 投機的実行 (CVE-2017-5715、CVE-2017-5753、CVE-2017-5754) に関連する脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。詳細については、「<https://aws.amazon.com/security/security-bulletins/AWS-2018-013/>」を参照してください。

## リリース 5.8.0

次のリリースノートには、Amazon EMR バージョン 5.8.0 リリースに関する情報が含まれています。Amazon EMR 5.7.0 リリースからの変更が含まれています。

初回リリース日: 2017 年 8 月 10 日

最新機能更新日: 2017 年 9 月 25 日

## アップグレード

このリリースでは、以下のアプリケーションおよびコンポーネントがアップグレードされ、以下のバージョンが含まれています。

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0。詳細については、Apache Hive サイトの「[リリースノート](#)」を参照してください。
- Spark 2.2.0。詳細については、Apache Spark サイトの「[リリースノート](#)」を参照してください。

## 新機能

- アプリケーション履歴の表示のサポートを追加しました (2017 年 9 月 25 日)。詳細については、「Amazon EMR 管理ガイド」の「[アプリケーションの履歴を表示する](#)」を参照してください。

## 変更点、機能強化、解決した問題

- AWS Glue Data Catalog との統合
  - Hive と Spark SQL が AWS Glue Data Catalog を Hive メタデータストアとして使用する機能を追加しました。詳細については、「[AWS Glue データカタログを Hive のメタストアとして使用](#)」

する」および「[AWS Glue データカタログを Spark SQL のメタストアとして使用する](#)」を参照してください。

- クラスター詳細の [Application history] を追加しました。これにより、YARN アプリケーションの履歴データや、Spark アプリケーションの追加の詳細を表示できます。詳細については、「Amazon EMR 管理ガイド」の「[アプリケーションの履歴を表示する](#)」を参照してください。
- Oozie
  - [OOZIE-2748](#) をバックポートしました。
- Hue
  - [HUE-5859](#) をバックポートしました。
- HBase
  - `getMasterInitializedTime` を使用して Java Management Extensions (JMX) 経由で HBase マスターサーバーの開始時間を公開するパッチを追加しました。
  - クラスターの開始時間を改善するパッチを追加しました。

## 既知の問題

- すべてのアプリケーションがインストールされ、デフォルトの Amazon EBS ルートボリュームサイズが変更されていない場合、クラスターの起動は失敗します。回避策として、の `aws emr create-cluster` コマンドを使用し AWS CLI、より大きな `--ebs-root-volume-size` パラメータを指定します。
- Hive 2.3.0 では `hive.compute.query.using.stats=true` がデフォルト設定になっています。これにより、クエリではデータが直接ではなく既存の統計から取得されるため、混乱が生じる場合があります。たとえば、`hive.compute.query.using.stats=true` が設定されたテーブルの LOCATION に新しいファイルをアップロードした場合、テーブルに対して `SELECT COUNT(*)` クエリを実行すると、追加された行がカウントされずに、統計からカウントが返されます。

回避策として、`ANALYZE TABLE` コマンドを使用して新しい統計を収集するか、`hive.compute.query.using.stats=false` を設定します。詳細については、Apache Hive ドキュメントの「[Statistics in Hive](#)」を参照してください。

- Spark - Spark を使用する場合、`appusher` デーモンには、長時間実行されている Spark ジョブで数時間または数日後に発生する可能性があるファイルハンドラのリークの問題があります。この問題を修正するには、マスターノードに接続し、「`sudo /etc/init.d/appusher stop`」と入力します。これにより、その `appusher` デーモンが停止し、Amazon EMR がこのデーモンを自動的に再起動します。

- アプリケーションの履歴
  - Spark のデッドエグゼキュターの履歴データは利用できません。
  - アプリケーション履歴は、セキュリティ設定を使用してインフライト暗号化を有効にするクラスターでは利用できません。

## リリース 5.7.0

次のリリースノートには、Amazon EMR 5.7.0 に関する情報が含まれています。Amazon EMR 5.6.0 リリースからの変更が含まれています。

リリース日: 2017 年 7 月 13 日

### アップグレード

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

### 新機能

- クラスターの作成時に、カスタム Amazon Linux AMI を指定する機能を追加しました。詳細については、「[カスタム AMI の使用](#)」を参照してください。

### 変更、機能強化、解決した問題

- HBase
  - HBase のリードレプリカクラスターを設定する機能を追加しました。詳細については、「[リードレプリカクラスターの使用](#)」を参照してください。
  - 複数のバグ修正と機能強化
- Presto - node.properties を設定する機能を追加しました。
- YARN - container-log4j.properties を設定する機能を追加しました。
- Sqoop - [SQOOP-2880](#) をバックポートしました。Sqoop 一時ディレクトリを設定できる引数が導入されています。

## リリース 5.6.0

次のリリースノートには、Amazon EMR 5.6.0 リリースに関する情報が含まれています。Amazon EMR 5.5.0 リリースからの変更が含まれています。

リリース日: 2017 年 6 月 5 日

### アップグレード

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0。Amazon EMR バージョン 5.0 以降の Spark 2.x をサポートする最初のバージョンの Mahout です。
- Spark 2.1.1

### 変更、機能強化、解決した問題

- Presto
  - セキュリティ設定を使って転送時の暗号化を有効にして、Presto ノード間で SSL/TLS を使った安全な通信を有効にする機能が追加されました。詳細については、「[転送時のデータ暗号化](#)」を参照してください。
  - [Presto 7661](#) を移植しました。これにより VERBOSE オプションを EXPLAIN ANALYZE ステートメントに追加し、クエリプランについての、より詳細なレポートと低レベルの統計を作成できます。

## リリース 5.5.3

次のリリースノートには、Amazon EMR リリース 5.5.3 に関する情報が含まれています。5.5.2 からの変更が含まれています。

初回リリース日: 2018 年 8 月 29 日

### 変更、機能強化、解決した問題

- このリリースでは、潜在的なセキュリティ脆弱性に対処しています。

## リリース 5.5.2

次のリリースノートには、Amazon EMR リリース 5.5.2 に関する情報が含まれています。5.5.1 からの変更が含まれています。

初回リリース日: 2018 年 3 月 29 日

変更、機能強化、解決した問題

- 潜在的な脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。

## リリース 5.5.1

次のリリースノートには、Amazon EMR 5.5.1 リリースに関する情報が含まれています。Amazon EMR 5.5.0 リリースからの変更が含まれています。

初回リリース日: 2018 年 1 月 22 日

変更、機能強化、解決した問題

- 投機的実行 (CVE-2017-5715、CVE-2017-5753、CVE-2017-5754) に関連する脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。詳細については、「<https://aws.amazon.com/security/security-bulletins/AWS-2018-013/>」を参照してください。

## リリース 5.5.0

次のリリースノートには、Amazon EMR 5.5.0 リリースに関する情報が含まれています。Amazon EMR 5.4.0 リリースからの変更が含まれています。

リリース日: 2017 年 4 月 26 日

アップグレード

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1

- ZooKeeper 3.4.10

## 変更点、機能強化、解決した問題

- Spark
  - Spark のパッチ「[\(SPARK-20115\) 外部シャッフルサービスが利用できない場合に、すべての失われたシャッフルブロックを再計算するように DAGScheduler を修正](#)」を Spark バージョン 2.1.0 にバックポートしました。これはこのリリースに含まれています。
- Flink
  - Flink は Scala 2.11 で作成されるようになりました。プロジェクトで Scala API とライブラリを使用する場合は、Scala 2.11 を使用することをお勧めします。
  - HADOOP\_CONF\_DIR と YARN\_CONF\_DIR のデフォルトが適切に設定されないため start-scala-shell.sh が機能しない問題に対応しました。さらに env.hadoop.conf.dir または env.yarn.conf.dir 設定分類の /etc/flink/conf/flink-conf.yaml と flink-conf を使って、これらの値を設定する機能を追加しました。
  - EMR 固有の新しいコマンドで、flink-scala-shell のラッパーとなる start-scala-shell.sh を追加しました。start-scala-shell に代えて、このコマンドを使用することをお勧めします。新しいコマンドにより実行が簡素化されます。たとえば、flink-scala-shell -n 2 は、タスクの並行度 2 で、Flink Scala シェルを開始します。
  - EMR 固有の新しいコマンドで、flink-yarn-session のラッパーとなる yarn-session.sh を追加しました。yarn-session に代えて、このコマンドを使用することをお勧めします。新しいコマンドにより実行が簡素化されます。たとえば、flink-yarn-session -d -n 2 は長時間稼働の Flink セッションを、デタッチ状態で、2 つのタスクマネージャを使って開始します。
  - 「[\(FLINK-6125\) Commons httpClient が Flink 1.2 でシェードされない](#)」の問題に対応しました。
- Presto
  - LDAP 認証のサポートが追加されました。Amazon EMR の Presto で LDAP を使用する場合は、Presto コーディネーターの HTTPS アクセスを有効にする必要があります (config.properties の http-server.https.enabled=true)。設定の詳細については、Presto のドキュメントの「[LDAP Authentication](#)」を参照してください。
  - SHOW GRANTS のサポートが追加されました。
- Amazon EMR の基本 Linux AMI
  - 現在、Amazon EMR リリースは Amazon Linux 2017.03 に基づいています。詳細については、「[Amazon Linux AMI 2017.03 リリースノート](#)」を参照してください。

- Python 2.6 は Amazon EMR の基本 Linux イメージから削除されました。Python 2.7 と 3.4 がデフォルトでインストールされます。必要な場合には Python 2.6 を手動でインストールできます。

## リリース 5.4.0

次のリリースノートには、Amazon EMR 5.4.0 リリースに関する情報が含まれています。Amazon EMR 5.3.0 リリースからの変更が含まれています。

リリース日: 2017 年 3 月 8 日

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- Flink 1.2.0 にアップグレードしました
- Hbase 1.3.0 にアップグレード済み
- Phoenix 4.9.0 にアップグレード済み

#### Note

古いバージョンの Amazon EMR から Amazon EMR バージョン 5.4.0 以降にアップグレードして、セカンダリインデックスを使用する場合、[Apache Phoenix のドキュメント](#)で説明されるようにローカルインデックスをアップグレードしてください。Amazon EMR は必要な設定を hbase-site 分類から削除しますが、インデックスはデータを再入力する必要があります。インデックスはオンラインとオフラインでアップグレードできます。オンラインのアップグレードがデフォルトです。これはバージョン4.8.0以降のPhoenixクライアントで初期化する間にインデックスの値が再設定されることを意味します。オフラインアップグレードを指定するには、`phoenix.client.localIndexUpgrade` 構成を `phoenix-site` 分類で `False` に設定してから、SSH をマスターノードに設定して `psql [zookeeper] -1`を実行します。

- Presto 0.166 にアップグレードしました
- Zeppelin 0.7.0 にアップグレードしました

## 変更と機能強化

以下は、リリースラベル emr-5.4.0 の Amazon EMR リリースでの変更点です。

- r4 インスタンスへのサポートを追加済み。「[Amazon EC2 インスタンスタイプ](#)」を参照してください。

## リリース 5.3.1

次のリリースノートには、Amazon EMR 5.3.1 リリースに関する情報が含まれています。Amazon EMR 5.3.0 リリースからの変更が含まれています。

リリース日: 2017 年 2 月 7 日

Zeppelin パッチをバックポートして Amazon EMR のデフォルトの AMI を更新するための小さな変更。

## リリース 5.3.0

次のリリースノートには、Amazon EMR 5.3.0 リリースに関する情報が含まれています。Amazon EMR 5.2.1 リリースからの変更が含まれています。

リリース日: 2017 年 1 月 26 日

## アップグレード

このリリースでは、次のアップグレードを使用できます。

- Hive 2.1.1 にアップグレードしました
- Hue 3.11.0 にアップグレードしました
- Spark 2.1.0 にアップグレードしました
- Oozie 4.3.0 にアップグレードしました
- Flink 1.1.4 にアップグレードしました

## 変更と機能強化

以下は、リリースラベル emr-5.3.0 の Amazon EMR リリースでの変更点です。



- `interpreters_shown_on_wheel` ファイルでの順序にかかわらず、ノートブックの選択ホイールで最初に表示するインタープリタを指定する `hue.ini` 設定を使用できるようにするパッチを Hue に追加しました。
- Hive の `hive-parquet-logging` ファイルで値を設定するために使用できる、`parquet-logging.properties` 設定分類を追加しました。

## リリース 5.2.2

次のリリースノートには、Amazon EMR 5.2.2 リリースに関する情報が含まれています。Amazon EMR 5.2.1 リリースからの変更が含まれています。

リリース日: 2017 年 5 月 2 日

### 以前のリリースから解決された既知の問題

- [SPARK-194459](#) をしました。char/varchar の列を持つ ORC テーブルからの読み取りが失敗する問題に対応しました。

## リリース 5.2.1

次のリリースノートには、Amazon EMR 5.2.1 リリースに関する情報が含まれています。Amazon EMR 5.2.0 リリースからの変更が含まれています。

リリース日: 2016 年 12 月 29 日

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- Presto を 0.157.1 にアップグレードしました。詳細については、Presto のドキュメントの「[Presto リリースノート](#)」を参照してください。
- ZooKeeper を 3.4.9 にアップグレードしました。詳細については、Apache ZooKeeper ドキュメントの「[リリースZooKeeper ノート](#)」を参照してください。

### 変更と機能強化

以下は、リリースラベル `emr-5.2.1` の Amazon EMR リリースでの変更点です。

- 5.0.0、5.0.3、5.2.0 を除く Amazon EMR バージョン 4.8.3 以降の Amazon EC2 の m4.16xlarge インスタンスタイプのサポートが追加されました。
- 現在、Amazon EMR リリースは Amazon Linux 2016.09 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/>」を参照してください。
- これで、Flink および YARN 設定パスの場所はデフォルトで /etc/default/flink に設定されましたので、Flink のジョブを起動するために FLINK\_CONF\_DIR または HADOOP\_CONF\_DIR ドライバースクリプトを実行するときに、環境変数 flink および yarn-session.sh を設定する必要はありません。
- FlinkKinesisConsumer クラスのサポートを追加しました。

## 以前のリリースから解決された既知の問題

- Hadoop で、レプリケーションと大きなクラスター内の同じファイルの削除の間に競合があるため、ReplicationMonitor スレッドが長時間停止する問題を修正しました。
- ジョブステータスが正常に更新されなかったときに ControlledJob、#toString が null ポインタ例外 (NPE) で失敗する問題を修正しました。

## リリース 5.2.0

次のリリースノートには、Amazon EMR 5.2.0 リリースに関する情報が含まれています。Amazon EMR 5.1.0 リリースからの変更が含まれています。

リリース日: 2016 年 11 月 21 日

### 変更と機能強化

このリリースでは、次の変更と機能強化を使用できます。

- HBase の Amazon S3 ストレージモードを追加しました。
- HBase ルートディレクトリの Amazon S3 ロケーションを指定できます。Amazon S3 の HBase の詳細については、「[Amazon S3 の HBase](#)」を参照してください。

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- Spark 2.0.2 にアップグレードしました

## 以前のリリースから解決された既知の問題

- EBS のみのインスタンスタイプで 2 TB に制約されていた /mnt に関する問題を修正。
- 通常の log4j-configured .log ファイルではなく、対応する .out ファイルに出力され、1 時間ごとにローテーションされていた、インスタンスコントローラーおよび logpusher ログに関する問題を修正。 .out ファイルはローテーションしないため、最終的には /emr パーティションがいっぱいになります。この問題は、ハードウェア仮想マシン (HVM) のインスタンスタイプにのみ影響します。

## リリース 5.1.0

次のリリースノートには、Amazon EMR 5.1.0 リリースに関する情報が含まれています。Amazon EMR 5.0.0 リリースからの変更が含まれています。

リリース日: 2016 年 11 月 3 日

### 変更と機能強化

このリリースでは、次の変更と機能強化を使用できます。

- Flink 1.1.3 のサポートを追加。
- Presto が、Hue のノートブックセクションでオプションとして追加。

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- HBase 1.2.3 にアップグレードしました
- Zeppelin 0.6.2 にアップグレードしました

## 以前のリリースから解決された既知の問題

- 以前の Amazon EMR 4.x バージョンと同じく、ORC ファイルがある Amazon S3 の Tez クエリが実行されない問題を修正しました。

## リリース 5.0.3

次のリリースノートには、Amazon EMR 5.0.3 リリースに関する情報が含まれています。Amazon EMR 5.0.0 リリースからの変更が含まれています。

リリース日: 2016 年 10 月 24 日

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- Hadoop 2.7.3 にアップグレードしました
- Presto 0.152.3 にアップグレードします。このアップグレードには Presto ウェブインターフェイスのサポートが含まれています。Presto コーディネーターの Presto ウェブインターフェイスには、ポート 8889 を使用してアクセスできます。Presto ウェブインターフェイスの詳細については、Presto のドキュメントの「[ウェブインターフェイス](#)」を参照してください。
- Spark 2.0.1 にアップグレードしました
- 現在、Amazon EMR リリースは Amazon Linux 2016.09 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/>」を参照してください。

## リリース 5.0.0

リリース日: 2016 年 7 月 27 日

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- Hive 2.1 にアップグレードしました
- Presto 0.150 にアップグレードしました
- Spark 2.0 にアップグレードしました
- Hue 3.10.0 にアップグレードしました
- Pig 0.16.0 にアップグレードしました
- Tez 0.8.4 にアップグレードしました
- Zeppelin 0.6.1 にアップグレードしました

## 変更と機能強化

以下は、リリースラベル emr-5.0.0 以降の Amazon EMR リリースでの変更点です。

- Amazon EMR は Hive (バージョン 2.1)、Pig (バージョン 0.16.0) の最新のオープンソースバージョンをサポートします。以前に Amazon EMR で Hive または Pig を使用している場合、このことはいくつかのユースケースに影響を与える可能性があります。詳細については、「[Hive](#)」および「[Pig](#)」を参照してください。
- 現在の Hive および Pig のデフォルト実行エンジンは Tez です。これを変更するには、hive-site および pig-properties のそれぞれの設定分類の適切な値を編集します。
- サービスが原因を識別できる場合にステップ障害の根本原因を表示できるようにするステップ、デバッグ機能が追加されました。詳細については、「Amazon EMR 管理ガイド」の「[デバッグの詳細ステップ](#)」を参照してください。
- アプリケーションは以前「-Sandbox」で終了しましたが、そのサフィックスはもうありません。たとえば、これらのアプリケーションを使ってクラスターを起動するスクリプトを使用している場合、これによってオートメーションが中断する可能性があります。次の表は、Amazon EMR 4.7.2 と Amazon EMR 5.0.0 のアプリケーション名を示しています。

### アプリケーション名変更

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-サンドボックス	ZooKeeper

- Spark は現在 Scala 2.11 向けにコンパイルされています。
- 現在のデフォルト JVM は Java 8 です。すべてのアプリケーションは Java 8 ランタイムを使用して動作します。アプリケーションのバイトコードターゲットには変更はありません。ほとんどのアプリケーションは、引き続き Java 7 を対象としています。
- Zeppelin には、認証機能が組み込まれています。詳細については、「[Zeppelin](#)」を参照してください。

- セキュリティ設定のサポートを追加しました。これにより、暗号化オプションをより簡単に作成、適用できます。詳細については、「[データの暗号化](#)」を参照してください。

## リリース 4.9.5

次のリリースノートには、Amazon EMR リリース 4.9.5 に関する情報が含まれています。4.9.4 からの変更が含まれています。

初回リリース日: 2018 年 8 月 29 日

変更、機能強化、解決した問題

- HBase
  - このリリースでは、潜在的なセキュリティ脆弱性に対処しています。

## リリース 4.9.4

次のリリースノートには、Amazon EMR リリース 4.9.4 に関する情報が含まれています。4.9.3 からの変更が含まれています。

初回リリース日: 2018 年 3 月 29 日

変更、機能強化、解決した問題

- 潜在的な脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しました。

## リリース 4.9.3

次のリリースノートには、Amazon EMR 4.9.3 リリースに関する情報が含まれています。Amazon EMR 4.9.2 リリースからの変更が含まれています。

初回リリース日: 2018 年 1 月 22 日

変更、機能強化、解決した問題

- 投機的実行 (CVE-2017-5715、CVE-2017-5753、CVE-2017-5754) に関連する脆弱性に対処するために、Amazon EMR のデフォルトの Amazon Linux AMI の Amazon Linux カーネルを更新しまし

た。詳細については、「<https://aws.amazon.com/security/security-bulletins/AWS-2018-013/>」を参照してください。

## リリース 4.9.2

次のリリースノートには、Amazon EMR 4.9.2 リリースに関する情報が含まれています。Amazon EMR 4.9.1 リリースからの変更が含まれています。

リリース日: 2017 年 7 月 13 日

このリリースでは小さな変更、バグ修正、および機能強化が行われました。

## リリース 4.9.1

次のリリースノートには、Amazon EMR 4.9.1 リリースに関する情報が含まれています。Amazon EMR 4.8.4 リリースからの変更が含まれています。

リリース日: 2017 年 4 月 10 日

### 以前のリリースから解決された既知の問題

- [HIVE-9976](#) および [HIVE-10106](#) を移植しました。
- 多数のノード (2,000 以上) やコンテナ (5,000 以上) によって、"Exception in thread 'main' java.lang.OutOfMemoryError" などのメモリエラーが発生することがある、YARN の問題を修正しました。

### 変更と機能強化

以下は、リリースラベル emr-4.9.1 の Amazon EMR リリースでの変更点です。

- 現在、Amazon EMR リリースは Amazon Linux 2017.03 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/>」を参照してください。
- Python 2.6 は Amazon EMR の基本 Linux イメージから削除されました。必要な場合には Python 2.6 を手動でインストールできます。

## リリース 4.8.4

次のリリースノートには、Amazon EMR 4.8.4 リリースに関する情報が含まれています。Amazon EMR 4.8.3 リリースからの変更が含まれています。

リリース日: 2017 年 2 月 7 日

このリリースでは小さな変更、バグ修正、および機能強化が行われました。

## リリース 4.8.3

次のリリースノートには、Amazon EMR 4.8.3 リリースに関する情報が含まれています。Amazon EMR 4.8.2 リリースからの変更が含まれています。

リリース日: 2016 年 12 月 29 日

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- Presto を 0.157.1 にアップグレードしました。詳細については、Presto のドキュメントの「[Presto リリースノート](#)」を参照してください。
- Spark を 1.6.3 にアップグレードしました。詳細については、Apache Spark のドキュメントの「[Spark リリースノート](#)」を参照してください。
- を ZooKeeper 3.4.9 にアップグレードしました。詳細については、Apache ZooKeeper ドキュメントの「[リリースZooKeeper ノート](#)」を参照してください。

### 変更と機能強化

以下は、リリースラベル emr-4.8.3 の Amazon EMR リリースでの変更点です。

- 5.0.0、5.0.3、5.2.0 を除く Amazon EMR バージョン 4.8.3 以降の Amazon EC2 の m4.16xlarge インスタンスタイプのサポートが追加されました。
- 現在、Amazon EMR リリースは Amazon Linux 2016.09 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/>」を参照してください。

### 以前のリリースから解決された既知の問題

- Hadoop で、レプリケーションと大きなクラスター内の同じファイルの削除の間に競合があるため、ReplicationMonitor スレッドが長時間停止する問題を修正しました。
- ジョブステータスが正常に更新されなかったときに ControlledJob、#toString が null ポインタ例外 (NPE) で失敗する問題を修正しました。



## リリース 4.8.2

次のリリースノートには、Amazon EMR 4.8.2 リリースに関する情報が含まれています。Amazon EMR 4.8.0 リリースからの変更が含まれています。

リリース日: 2016 年 10 月 24 日

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- Hadoop 2.7.3 にアップグレードしました
- Presto 0.152.3 にアップグレードします。このアップグレードには Presto ウェブインターフェイスのサポートが含まれています。Presto コーディネーターの Presto ウェブインターフェイスには、ポート 8889 を使用してアクセスできます。Presto ウェブインターフェイスの詳細については、Presto のドキュメントの「[ウェブインターフェイス](#)」を参照してください。
- 現在、Amazon EMR リリースは Amazon Linux 2016.09 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/>」を参照してください。

## リリース 4.8.0

リリース日: 2016 年 9 月 7 日

### アップグレード

このリリースでは、次のアップグレードを使用できます。

- HBase 1.2.2 にアップグレードしました
- Presto-Sandbox 0.151 にアップグレードしました
- Tez 0.8.4 にアップグレードしました
- Zeppelin-Sandbox 0.6.1 にアップグレードしました

### 変更と機能強化

以下は、リリースラベル emr-4.8.0 の Amazon EMR リリースでの変更点です。

- YARN で、インスタンスが終了したために がもう存在しないコンテナをクリーンアップ ApplicationMaster しようとする問題を修正しました。

- Oozie の例の Hive2 アクションの hive-server2 URL を修正しました。
- さらに多くの Presto カタログのサポートを追加しました。
- 次のパッチを移植しました:  
[HIVE-8948](#)、[HIVE-12679](#)、[HIVE-13405](#)、[PHOENIX-3116](#)、[HADOOP-12689](#)
- セキュリティ設定のサポートを追加しました。これにより、暗号化オプションをより簡単に作成、適用できます。詳細については、「[データの暗号化](#)」を参照してください。

## リリース 4.7.2

以下のリリースノートには、Amazon EMR 4.7.2 に関する情報が含まれています。

リリース日: 2016 年 7 月 15 日

### 機能

このリリースでは、次の機能を使用できます。

- Mahout 0.12.2 にアップグレードしました
- Presto 0.148 にアップグレードしました
- Spark 1.6.2 にアップグレードしました
- パラメータとして URI を使用して、EMRFS AWSCredentialsProvider で使用する を作成できるようになりました。詳細については、「[EMRFS 用の を作成する AWSCredentialsProvider](#)」を参照してください。
- EMRFS では、ユーザーが emrfs-site.xml で fs.s3.consistent.dynamodb.endpoint プロパティを使用して、整合性のあるビューのメタデータのカスタム DynamoDB エンドポイントを設定できるようになりました。
- /usr/bin に spark-example というスクリプトを追加しました。これにより /usr/lib/spark/spark/bin/run-example をラップし、例を直接実行できます。例えば、Spark ディストリビューションに付属する SparkPi 例を実行するには、コマンドライン spark-example SparkPi 100から を実行するか、API のステップ command-runner.jarとして を使用します。

### 以前のリリースから解決された既知の問題

- Oozie で、Spark もインストールされたときに spark-assembly.jar が正しい場所がないために、Oozie で Spark アプリケーションを起動できなくなる問題を解決しました。

- YARN コンテナで Spark Log4j ベースのログインに関する問題を修正しました。

## リリース 4.7.1

リリース日: 2016 年 6 月 10 日

### 以前のリリースから解決された既知の問題

- プライベートサブネットを持つ VPC で起動されたクラスターの起動時間が長くなる問題を修正しました。このバグの影響があったのは、Amazon EMR 4.7.0 リリースで起動されたクラスターのみです。
- Amazon EMR 4.7.0 リリースで起動されたクラスターに対する Amazon EMR でのファイルのリスト処理が不適切であった問題を修正しました。

## リリース 4.7.0

### Important

Amazon EMR 4.7.0 は廃止されています。代わりに Amazon EMR 4.7.1 以降を使用してください。

リリース日: 2016 年 6 月 2 日

### 機能

このリリースでは、次の機能を使用できます。

- Apache Phoenix 4.7.0 を追加しました
- Apache Tez 0.8.3 を追加しました
- HBase 1.2.1 にアップグレードしました
- Mahout 0.12.0 にアップグレードしました
- Presto 0.147 にアップグレードしました
- AWS SDK for Java を 1.10.75 にアップグレードしました
- ユーザーがローカルモードで Pig を実行できるようにするため、最終的なフラグが `mapreduce.cluster.local.dir` の `mapred-site.xml` プロパティから削除されました。

## Amazon Redshift JDBC ドライバーがクラスターで利用可能

Amazon Redshift JDBC ドライバーが `/usr/share/aws/redshift/jdbc` に含まれました。`/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar` は JDBC 4.1 互換の Amazon Redshift ドライバー、`/usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar` は JDBC 4.0 互換の Amazon Redshift ドライバーです。詳細については、「Amazon Redshift 管理ガイド」の「[JDBC 接続の設定](#)」を参照してください。

## Java 8

Presto を除き、OpenJDK 1.7 はすべてのアプリケーションに使用されるデフォルトの JDK です。ただし、OpenJDK 1.7 と 1.8 の両方がインストールされています。アプリケーションの `JAVA_HOME` を設定する方法については、「[Java 8 を使用したアプリケーションの設定](#)」を参照してください。

## 以前のリリースから解決された既知の問題

- `emr-4.6.0` で Amazon EMR 用のスループット最適化 HDD (st1) EBS ボリュームで著しくパフォーマンスに影響を与えていたカーネルの問題を修正しました。
- アプリケーションとして Hadoop を選択せずに HDFS 暗号化ゾーンを指定した場合にクラスターが失敗する問題を修正しました。
- デフォルトの HDFS 書き込みポリシーを RoundRobin から AvailableSpaceVolumeChoosingPolicy に変更しました。一部のボリュームが RoundRobin 設定で適切に使用されなかったため、コアノードに障害が発生し、HDFS の信頼性が低下しました。
- 整合性のあるビューを実現するためにデフォルトの DynamoDB メタデータテーブルを作成するときに、例外が発生する原因となっていた EMRFS CLI に関する問題を修正しました。
- マルチパートの名前の変更およびコピーオペレーション中に発生する可能性のあった、EMRFS のデッドロックの問題を修正しました。
- EMRFS で CopyPart サイズがデフォルト 5 MB になる問題を修正しました。現在では、デフォルト値は 128 MB で適切に設定されます。
- サービスを停止できなくなる可能性のある、Zeppelin upstart 設定の問題を修正しました。
- `s3a://` がそれぞれのクラスパスで適切にロードされていないために、`/usr/lib/hadoop/hadoop-aws.jar` URI スキームを使用できなくなる Spark および Zeppelin の問題を修正しました。
- [HUE-2484](#) を移植しました。

- HBase ブラウザサンプルでの問題を修正するため、Hue 3.9.0 (JIRA が存在しない) から [commit](#) を移植しました。
- [HIVE-9073](#) を移植しました。

## リリース 4.6.0

リリース日: 2016 年 4 月 21 日

### 機能

このリリースでは、次の機能を使用できます。

- HBase 1.2.0 を追加しました
- Zookeeper-Sandbox 3.4.8 を追加しました
- Presto-Sandbox 0.143 にアップグレードしました
- 現在、Amazon EMR リリースは Amazon Linux 2016.03.0 に基づいています。詳細については、「<https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/>」を参照してください。

### スループット最適化 HDD (st1) EBS ボリュームタイプに影響を及ぼす問題

Linux カーネルバージョン 4.2 以降の問題は、EMR 用のスループット最適化 HDD (st1) EBS ボリュームのパフォーマンスに大きな影響を及ぼします。このリリース (emr-4.6.0) ではカーネルバージョン 4.4.5 を使用するため、影響を受けます。したがって、st1 EBS ボリュームを使用する場合、emr-4.6.0 を使用しないことをお勧めします。emr-4.5.0 以前の Amazon EMR リリースと st1 であれば、影響を受けずに使用できます。これに加えて、将来のリリースで修正が提供されます。

### Python のデフォルト値

現在、Python 3.4 がデフォルトでインストールされますが、Python 2.7 はシステムデフォルトのままです。いずれかのブートストラップアクションを使用して、Python 3.4 をシステムのデフォルトとして設定できます。設定 API を使用して、が使用する Python バージョンに影響を与えるように `spark-env`、`分類/usr/bin/python3.4` で `PYSPARK_PYTHON` エクスポートを に設定することができます PySpark。

## Java 8

Presto を除き、OpenJDK 1.7 はすべてのアプリケーションに使用されるデフォルトの JDK です。ただし、OpenJDK 1.7 と 1.8 の両方がインストールされています。アプリケーションの JAVA\_HOME を設定する方法については、「[Java 8 を使用したアプリケーションの設定](#)」を参照してください。

### 以前のリリースから解決された既知の問題

- アプリケーションのプロビジョニングが、生成されたパスワードが原因でランダムに失敗する問題を修正しました。
- 以前は、mysqld がすべてのノードにインストールされました。現在では、選択されたアプリケーションにコンポーネントとして mysql-server が含まれている場合のみ、マスターインスタンスのみにインストールされます。現在、HCatalog、Hive、Hue、Presto-Sandbox、および Sqoop-Sandbox の各アプリケーションに、mysql-server コンポーネントが含まれています。
- yarn.scheduler.maximum-allocation-vcores をデフォルトの 32 から 80 に変更しました。これにより、コアインスタスタイプが、YARN vcores が 32 より高く設定されているいくつかのラージインスタスタイプのいずれかであるクラスターで、maximizeResourceAllocation オプションを使用中に Spark で主に発生する、emr-4.4.0 での問題が修正されました。この問題の影響を受けていたのは、c4.8xlarge、cc2.8xlarge、hs1.8xlarge、i2.8xlarge、m2.4xlarge、r3.8xlarge、d2.8xlarge、または m4.10xlarge です。
- 現在では s3-dist-cp はすべての Amazon S3 候補に EMRFS を使用し、一時 HDFS ディレクトリにはステージングしません。
- クライアント側の暗号化のマルチパートアップロードの例外処理に関する問題を修正しました。
- ユーザーが Amazon S3 ストレージクラスを変更できるようにするオプションを追加しました。デフォルトでは、この設定は STANDARD です。emrfs-site 設定の分類設定は fs.s3.storageClass で、指定できる値は STANDARD、STANDARD\_IA、REDUCED\_REDUNDANCY です。ストレージクラスの詳細については、「Amazon Simple Storage Service ユーザーガイド」の「[ストレージクラス](#)」を参照してください。

## リリース 4.5.0

リリース日: 2016 年 4 月 4 日

## 機能

このリリースでは、次の機能を使用できます。

- Spark 1.6.1 にアップグレードしました
- Hadoop 2.7.2 にアップグレードしました
- Presto 0.140 にアップグレードしました
- Amazon S3 サーバー側の暗号化 AWS KMS のサポートが追加されました。

## 以前のリリースから解決された既知の問題

- ノードが再起動された後に MySQL および Apache サーバーが起動しない問題を修正しました。
- Amazon S3 に保存されているパーティション分割されていないテーブルで IMPORT が正しく機能しない問題を修正しました。
- Hive テーブルに書き込むときに、ステージングディレクトリが /mnt/tmp ではなく /tmp であることが要求される Presto の問題を修正しました。

## リリース 4.4.0

リリース日: 2016 年 3 月 14 日

## 機能

このリリースでは、次の機能を使用できます。

- HCatalog 1.0.0 を追加しました
- Sqoop-Sandbox 1.4.6 を追加しました
- Presto 0.136 にアップグレードしました
- Zeppelin 0.5.6 にアップグレードしました
- Mahout 0.11.1 にアップグレードしました
- デフォルトで `dynamicResourceAllocation` を有効にしました。
- リリースのすべての設定分類の表を追加しました。詳細については、「[アプリケーションの設定](#)」の設定分類の表を参照してください。

## 以前のリリースから解決された既知の問題

- `maximizeResourceAllocation` 設定で YARN ApplicationMaster デーモンに十分なメモリを予約できない問題を修正しました。
- カスタム DNS で発生した問題を修正しました。 `resolve.conf` のエントリが、提供されたカスタムエントリよりも前に指定されている場合、そのカスタムエントリは解決されません。この動作は、デフォルトの VPC ネームサーバーが `resolve.conf` のトップエントリとして挿入された VPC のクラスターによって影響を受けました。
- デフォルトの Python がバージョン 2.7 に移行した場合に、そのバージョンに対して boto がインストールされなかった問題を修正しました。
- YARN コンテナと Spark アプリケーションが独自の Ganglia ラウンドロビンデータベース (rrd) ファイルを生成し、それによりインスタンスにアタッチされた最初のディスクがいっぱいになる問題を修正しました。この修正によって、YARN コンテナレベルのメトリクスが無効になり、Spark アプリケーションレベルのメトリクスが無効になりました。
- ログプッシャーですべての空のログフォルダーが削除される問題を修正しました。この問題により、ログプッシャーは `user` の空の `/var/log/hive` フォルダを削除したため、Hive CLI はログを記録できませんでした。
- パーティション分割に影響し、インポート中にエラーを発生させた、Hive のインポートに影響を与える問題を修正しました。
- EMRFS と `s3-dist-cp` が、ピリオドを含むバケット名を適切に処理しなかった問題を修正しました。
- EMRFS の動作を変更し、バージョンングが有効なバケットで、`_${folder}$` マーカーファイルが連続して作成されないようにしました。これにより、バージョンングが有効なバケットでパフォーマンスが向上する可能性があります。
- クライアント側の暗号化が有効になっている場合を除き、インストラクションファイルを使用しないよう EMRFS の動作を変更しました。クライアント側の暗号化を使用中にインストラクションファイルを削除する場合は、`emrfs-site.xml` プロパティの `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled` を `true` に設定できます。
- YARN ログの集計を変更し、集計先でログを 2 日間保持するようにしました。デフォルトの送信先はクラスターの HDFS ストレージです。この期間を変更する場合は、クラスターの作成時に `yarn.log-aggregation.retain-seconds` 設定分類を使用して `yarn-site` の値を変更します。これまでどおり、クラスターの作成時は、`log-uri` パラメータを使用して、Amazon S3 にアプリケーションログを保存できます。



## 適用されたパッチ

オープンソースのプロジェクトから、次のパッチがこのリリースで追加されました。

- [HIVE-9655](#)
- [HIVE-9183](#)
- [HADOOP-12810](#)

## リリース 4.3.0

リリース日: 2016 年 1 月 19 日

### 機能

このリリースでは、次の機能を使用できます。

- Hadoop 2.7.1 にアップグレードしました
- Spark 1.6.0 にアップグレードしました
- Ganglia を 3.7.2 にアップグレードしました
- Presto を 0.130 にアップグレードしました

Amazon EMR を true に設定すると、`spark.dynamicAllocation.enabled` にくつつかの変更が加えられていました。デフォルトでは false です。true に設定すると、`maximizeResourceAllocation` 設定で定義されているデフォルト設定に影響を与えます。

- `spark.dynamicAllocation.enabled` を true に設定した場合、`spark.executor.instances` は `maximizeResourceAllocation` によって設定されません。
- `spark.driver.memory` 設定は、`spark.executors.memory` 設定と同様に、クラスター内のインスタンスタイプに基づいて定義されます。ただし、Spark ドライバーアプリケーションは、マスターインスタンスまたはいずれかのコアインスタンスで (たとえば、YARN クライアントモードとクラスターモードのそれぞれで) 実行されるため、`spark.driver.memory` 設定は、これらの 2 つのインスタンスグループ間で、小さい方のインスタンスのインスタンスタイプに基づいて定義されます。
- `spark.default.parallelism` 設定は、YARN コンテナに使用可能な CPU コアの数の 2 倍に定義されます。以前のリリースでは、半分の値に定義されていました。

- Spark YARN プロセス用に予約されるメモリアーヘッドの計算精度が上がったため、Spark に使用可能なメモリの合計量 ( `spark.executor.memory` ) がわずかに増えました。

## 以前のリリースから解決された既知の問題

- 現在、YARN ログの集計はデフォルトで有効になります。
- YARN ログの集計が有効な場合に、クラスターの Amazon S3 ログバケットにログがプッシュされない問題を修正しました。
- YARN コンテナサイズは、すべてのノードタイプで新たに最低 32 になりました。
- 大規模なクラスターのマスターノードで過剰なディスク I/O を発生させる Ganglia の問題を修正しました。
- クラスターのシャットダウン時に Amazon S3 にアプリケーションログがプッシュされない問題を修正しました。
- 特定のコマンドを失敗させる EMRFS CLI の問題を修正しました。
- Zeppelin で、基盤となる に依存関係がロードされない問題を修正しました SparkContext。
- インスタンスの追加を試みるサイズ変更の発行によって発生する問題を修正しました。
- CREATE TABLE AS SELECT が Amazon S3 への過剰なリスト呼び出しを行う Hive の問題を修正しました。
- Hue、Oozie、および Ganglia がインストールされていると、大規模なクラスターが適切にプロビジョニングされない問題を修正しました。
- エラーで失敗した場合でもゼロ終了コードを返す s3-dist-cp の問題を修正しました。

## 適用されたパッチ

オープンソースのプロジェクトから、次のパッチがこのリリースで追加されました。

- [OOZIE-2402](#)
- [HIVE-12502](#)
- [HIVE-10631](#)
- [HIVE-12213](#)
- [HIVE-10559](#)
- [HIVE-12715](#)
- [HIVE-10685](#)

## リリース 4.2.0

リリース日: 2015 年 11 月 18 日

### 機能

このリリースでは、次の機能を使用できます。

- Ganglia のサポートを追加しました
- Spark 1.5.2 にアップグレードしました
- Presto 0.125 にアップグレードしました
- Oozie を 4.2.0 にアップグレードしました
- Zeppelin を 0.5.5 にアップグレードしました
- を 1.10.27 AWS SDK for Java にアップグレードしました

### 以前のリリースから解決された既知の問題

- デフォルトのメタデータテーブル名を使用しない EMRFS CLI の問題を修正しました。
- Amazon S3 で ORC-backed テーブルを使用するときに発生した問題を修正しました。
- Spark 設定で Python バージョンが一致しない問題を修正しました。
- VPC のクラスターでの DNS の問題により、YARN ノードのステータスが報告されない問題を修正しました。
- YARN がノードを廃棄することが原因でアプリケーションがハングしたり、新しいアプリケーションを予定できなくなったりする問題を修正しました。
- クラスターが TIMED\_OUT\_STARTING というステータスで終了するときに発生する問題を修正しました。
- EMRFS Scala 依存関係を他のビルドに含める場合に発生する問題を修正しました。Scala 依存関係が削除されました。

# アプリケーションの設定

アプリケーションのデフォルトの設定を上書きするために、設定オブジェクトを指定できます。短縮構文を使用して設定を指定するか、JSON ファイルの設定オブジェクトを参照することができます。設定オブジェクトは、分類、プロパティ、オプションの入れ子になっている設定で構成されます。プロパティは、変更するアプリケーション設定に対応します。単一の JSON オブジェクトで、複数のアプリケーションに複数の分類を指定できます。

## Warning

Amazon EMR の Describe API オペレーションと List API オペレーションにより、Amazon EMR ジョブフローの一部として使用するカスタム設定と構成可能な設定がプレーンテキストで出力されます。こうした設定でパスワードなどの機密情報を指定する方法については、「AWS Secrets Manager」の「[機密の設定データを保存する](#)」を参照してください。

Amazon EMR リリースバージョンによって使用可能な設定分類は異なります。特定のリリースバージョンでサポートされている設定分類のリストについては、「[Amazon EMR リリースについて](#)」にあるそのリリースバージョンのページを参照してください。

設定の一覧の JSON ファイル例を次に示します。

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.security.groups.cache.secs": "250"
    }
  },
  {
    "Classification": "mapred-site",
    "Properties": {
      "mapred.tasktracker.map.tasks.maximum": "2",
      "mapreduce.map.sort.spill.percent": "0.90",
      "mapreduce.tasktracker.reduce.tasks.maximum": "5"
    }
  }
]
```

設定分類は、多くの場合アプリケーション固有の設定ファイルにマップします。たとえば、hive-site 分類は Hive の hive-site.xml 設定ファイルの設定にマッピングされます。この例外には、--namenode-heap-size のような環境パラメータの設定に使用される、サポートされなくなったブートストラップアクション configure-daemons があります。このようなオプションはそれ自身の下位エクスポート分類とともに、hadoop-env および yarn-env 分類の中に入れ子になっています。env で終わる分類には、下位エクスポート分類を使用してください。

もう 1 つの例外として s3get があり、これはクライアント側の暗号化で使用するために、クラスター内の各ノードに顧客の EncryptionMaterialsProvider オブジェクトを配置するのに使用されます。この用途には、emrfs-site 分類にオプションが追加されています。

hadoop-env 分類の例を次に示します。

```
[
  {
    "Classification": "hadoop-env",
    "Properties": {

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "HADOOP_DATANODE_HEAPSIZE": "2048",
          "HADOOP_NAMENODE_OPTS": "-XX:GCTimeRatio=19"
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

yarn-env 分類の例を次に示します。

```
[
  {
    "Classification": "yarn-env",
    "Properties": {
```

```

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "YARN_RESOURCEMANAGER_OPTS": "-Xdebug -Xrunjdp:transport=dt_socket"
        },
        "Configurations": [
          ]
        ]
      }
    ]
  }
}
]

```

次の設定は設定ファイルに属しませんが、ユーザーに代わって複数の設定を構成できるようにするために Amazon EMR が使用します。

#### Amazon EMR-によって選別された設定

アプリケーション	リリースラベルの分類	有効なプロパティ	どのようなときに使うか
Spark	spark	maximizeResourceAllocation	各ノードの最大リソースを利用するようにエグゼキューターを設定します。

#### トピック

- [クラスターの作成時のアプリケーションの設定](#)
- [実行中のクラスター内のインスタンスグループの再設定](#)
- [機密設定データを AWS Secrets Manager に保存する](#)
- [特定の Java 仮想マシンを使用するようにアプリケーションを設定](#)

## クラスターの作成時のアプリケーションの設定

クラスターを作成するときは、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用して、アプリケーションのデフォルト設定を上書きできます。

アプリケーションのデフォルト設定をオーバーライドするには、設定分類でカスタム値を指定します。設定分類は、アプリケーションを設定する XML ファイル (hive-site.xml など) に対応しています。

設定分類は、Amazon EMR のリリースバージョンによって異なります。特定のリリースバージョンで利用可能な設定分類のリストについては、「[Amazon EMR リリース 6.4.0](#)」などのリリース詳細ページを参照してください。

## クラスター作成時にコンソールで設定を指定する

設定を指定するには、クラスターの作成ページに移動し、ソフトウェア設定 を展開します。コンソールで JSON またはシャドウテキストに倣った短縮構文を使用して、設定を直接入力できます。それ以外の方法として、JSON の Configurations オブジェクトとしてファイルの Amazon S3 URI を指定することもできます。

インスタンスグループの設定を指定するには、クラスターのリストでクラスターを選択し、設定タブを選択します。インスタンスグループ設定 テーブルで、編集するインスタンスグループを選択し、の再設定 を選択します。

## クラスターの作成 AWS CLI 時に を使用して設定を指定する

ローカルまたは Amazon S3 に保存された JSON ファイルのパスを指定することで、create-cluster に対して設定を指定できます。次の例は、Amazon EMR にデフォルトのロールを使用していて、そのロールが作成されていることを前提としています。ロールを作成する必要がある場合は、最初に aws emr create-default-roles を実行します。

設定がローカルディレクトリにある場合は、次のコマンド例を使用できます。

```
aws emr create-cluster --use-default-roles --release-label emr-7.1.0 --applications
  Name=Hive \  
--instance-type m5.xlarge --instance-count 3 --configurations file:///./
configurations.json
```

設定が Amazon S3 パスに含まれている場合は、次の次善策を設定した後に、Amazon S3 create-cluster パスをコマンドに渡す必要があります。

```
#!/bin/sh
# Assume the ConfigurationS3Path is not public, and its present in the same AWS account
as the EMR cluster
ConfigurationS3Path="s3://my-bucket/config.json"
```

```
# Get a presigned HTTP URL for the s3Path
ConfigurationURL=`aws s3 presign $ConfigurationS3Path --expires-in 300`
# Fetch the presigned URL, and minify the JSON so that it spans only a single line
Configurations=`curl $ConfigurationURL | jq -c .`
aws emr create-cluster --use-default-roles --release-label emr-5.34.0 --instance-type
m5.xlarge --instance-count 2 --applications Name=Hadoop Name=Spark --configurations
$Configurations
```

## クラスター作成時に Java SDK を使用して設定を指定する

AWS SDK for Javaを使用して設定を指定する方法を次のプログラム抜粋に示します。

```
Application hive = new Application().withName("Hive");

Map<String,String> hiveProperties = new HashMap<String,String>();
hiveProperties.put("hive.join.emit.interval","1000");
hiveProperties.put("hive.merge.mapfiles","true");

Configuration myHiveConfig = new Configuration()
    .withClassification("hive-site")
    .withProperties(hiveProperties);

RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("Create cluster with ReleaseLabel")
    .withReleaseLabel("emr-5.20.0")
    .withApplications(hive)
    .withConfigurations(myHiveConfig)
    .withServiceRole("EMR_DefaultRole")
    .withJobFlowRole("EMR_EC2_DefaultRole")
    .withInstances(new JobFlowInstancesConfig()
        .withEc2KeyName("myEc2Key")
        .withInstanceCount(3)
        .withKeepJobFlowAliveWhenNoSteps(true)
        .withMasterInstanceType("m4.large")
        .withSlaveInstanceType("m4.large")
    );
```

## 実行中のクラスター内のインスタンスグループの再設定

Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスターアプリケーションを再設定して追加の設定分類を指定できます。そのためには、Amazon



EMR コンソール、AWS Command Line Interface ( AWS CLI )、または SDK AWS を使用できません。

新しい Amazon EMR コンソールでインスタンスグループのアプリケーション設定を更新すると、新規設定と既存の設定がマージされ、有効な新規設定が作成されます。Amazon EMR が設定をマージできないという稀なケースでは、コンソールに警告が表示されます。

インスタンスグループの再設定リクエストを送信すると、Amazon EMR が新しい設定仕様にバージョン番号を割り当てます。CloudWatch イベントを表示することで、設定のバージョン番号またはインスタンスグループの状態を追跡できます。詳細については、「[イベントのモニタリング CloudWatch](#)」を参照してください。

#### Note

クラスターの作成時に指定したクラスター設定に対して実行できるのは上書きのみであり、削除はできません。既存の設定と指定したファイルに差異がある場合、Amazon EMR は SSH を使用してクラスターに接続中に変更した設定などの手動で変更された設定を、指定したインスタンスグループのクラスターのデフォルト値にリセットします。

## インスタンスグループの再設定時の考慮事項

### 再設定アクション

Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用して再設定リクエストを送信すると、Amazon EMR は既存のクラスター上の設定ファイルをチェックします。既存の設定と指定したファイルに差異がある場合、Amazon EMR は再設定アクションを開始し、いくつかのアプリケーションを再起動し、SSH を使用してクラスターに接続中に変更した設定などの手動で変更された設定を、指定したインスタンスグループのクラスターのデフォルト値にリセットします。

#### Note

Amazon EMR は、インスタンスグループの再設定ごとに、デフォルトのアクションをいくつか実行します。これらのデフォルトのアクションが、ユーザーが行ったクラスターのカスタマイズと競合して、再設定に失敗することがあります。再設定失敗のトラブルシューティング方法については、[インスタンスグループの再設定のトラブルシューティング](#)を参照してください。

Amazon EMR は、リクエストで指定した設定分類に対する再設定アクションも開始します。これらのアクションの詳細なリストについては、使用している Amazon EMR のバージョンの「設定分類」セクションを参照してください。例えば、[6.2.0 の設定分類](#)です。

#### Note

Amazon EMR リリースガイドには、Amazon EMR バージョン 5.32.0 および 6.2.0 以降の再設定アクションのみが記載されています。

## サービスの中断

Amazon EMR では、ローリングプロセスに従って、タスクインスタンスグループとコアインスタンスグループのインスタンスの再設定を行います。インスタンスグループ内のインスタンスのうち、一度に変更および再起動されるのはわずか 10% です。このプロセスは完了するまでに時間がかかりますが、実行中のクラスターでアプリケーションが失敗する可能性は低くなります。

YARN の再起動中に YARN ジョブを実行するには、複数のマスターノードを持つ Amazon EMR クラスターを作成するか、`yarn-site` 設定分類で `yarn.resourcemanager.recovery.enabled` を `true` に設定します。複数のマスターノードの使用の詳細については、「[高可用性 YARN ResourceManager](#)」を参照してください。

## アプリケーションの検証

Amazon EMR は、再設定の再起動プロセスの後に、クラスター上の各アプリケーションが実行中であることを確認します。使用できないアプリケーションがある場合、再設定操作全体が失敗します。再設定操作に失敗した場合、Amazon EMR は、設定パラメータを前に機能していたバージョンに戻します。


#### Note

再設定の失敗を回避するため、クラスターに使用する予定のアプリケーションのみをインストールすることをお勧めします。再設定リクエストを送信する前に、すべてのクラスターアプリケーションが正常に稼働していることを確認することもお勧めします。

## 再設定のタイプ

インスタンスグループは、2 つの方法のいずれかで再設定できます。

- 上書き: デフォルトの再設定方法です。5.35.0 および 6.6.0 より前の Amazon EMR リリースでは、この方法のみ使用できます。この方法で再設定すると、クラスター上のすべてのファイルが、新しく送信した設定セットで強制的に上書きされます。この方法では、再設定 API 以外で行った設定ファイル変更がすべて消去されます。
- マージ: Amazon EMR リリース 5.35.0 および 6.6.0 以降でサポートされる再設定の方法です。ただし、バージョンに関係なく、Amazon EMR コンソールでは、この再設定は行えません。この方法で再設定すると、新しく送信した設定が、クラスターに既に存在する設定と統合されます。このオプションでは、送信した新しい設定のみが追加または変更されます。既存の設定は保持されます。

 Note

Amazon EMR では、適切なサービス稼働に必要な Hadoop 設定の一部が継続的に上書きされます。

## 制約事項

実行中のクラスターのインスタンスグループを再設定するときは、次の制限を考慮してください。

- 非 Yarn アプリケーションでは、アプリケーションが適切に設定されていない場合は特に、再起動中の失敗やクラスターの問題が発生する可能性があります。メモリや CPU 使用率が最大に近づいているクラスターは、再起動プロセス後に問題が発生する可能性があります。これは、マスターインスタンスグループに対して特にあてはまります。
- インスタンスグループのサイズ変更中は、再設定リクエストを送信できません。インスタンスグループのサイズ変更中に再設定を開始した場合、インスタンスグループのサイズ変更が完了するまで再設定を開始することはできません。再設定中にサイズ変更を行う場合も同様です。
- インスタンスグループを再設定後、Amazon EMR は、新しい設定を反映するためにアプリケーションを再起動します。再設定中にアプリケーションが使用されていると、ジョブが失敗したり、アプリケーションで予期しない動作が発生したりする可能性があります。
- インスタンスグループの再設定に失敗した場合、Amazon EMR は、設定パラメータを前に機能していたバージョンに戻します。バージョンを戻せなかった場合は、新しい `ModifyInstanceGroup` リクエストを送信して、インスタンスグループを `SUSPENDED` 状態から復旧させる必要があります。
- Phoenix 設定分類の再設定リクエストは、Amazon EMR バージョン 5.23.0 以降でのみサポートされていて、Amazon EMR バージョン 5.21.0 や 5.22.0 ではサポートされていません。

- HBase 設定分類の再設定リクエストは、Amazon EMR バージョン 5.30.0 以降でのみサポートされていて、Amazon EMR バージョン 5.23.0 から 5.29.0 ではサポートされていません。
- Amazon EMR バージョン 5.27.0 以降でのみ、複数のプライマリノードを持つ EMR クラスターでのアプリケーション再設定リクエストがサポートされています。
- `hdfs-encryption-zones` 分類または Hadoop KMS 設定分類の再設定は、複数のプライマリノードを持つ Amazon EMR クラスターではサポートされていません。
- Amazon EMR は現在、YARN の再起動を必要とするキャパシテイスケジューラの特定の再設定リクエストをサポートしていません ResourceManager。例えば、キューを完全に削除することはできません。

## コンソールでのインスタンスグループの再設定

### Note

Amazon EMR コンソールでは、マージタイプの再設定がサポートされていません。

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [名前] の下のクラスター一覧で、再設定するアクティブなクラスターを選択します。
3. クラスターのクラスター詳細ページを開き、[設定] タブに移動します。
4. [フィルタ] ドロップダウンリストで、再設定するインスタンスグループを選択します。
5. [再設定] ドロップダウンメニューで、[テーブルで編集する] と [JSON ファイルで編集する] のいずれかを選択します。
  - テーブルで編集する -- 設定分類テーブルで、既存の設定のプロパティと値を編集するか、[設定の追加] を選択して追加の設定分類を指定します。
  - JSON ファイルで編集する -- JSON に直接設定を入力するか、短縮構文 (シャドウテキストで示されている) を使用します。それ以外の方法として、JSON の Configurations オブジェクトとしてファイルの Amazon S3 URI を指定します。

### Note

設定分類テーブルの [ソース] 列は、クラスター作成時と、このインスタンスグループ用の追加設定の指定時のどちらで設定を指定するかを示しています。インスタンスグルー

プの設定は、両方のソースから編集することができます。最初のクラスター設定を削除することはできませんが、インスタンスグループの設定は上書きすることができます。また、ネストされた設定分類は、テーブルで直接追加または編集することもできます。たとえば、`export` の追加の `hadoop-env` サブ分類を指定するには、テーブルで `hadoop.export` 設定分類を追加します。次に、この分類の特定のプロパティおよび値を指定します。

- (オプション) [Apply this configuration to all active instance groups (この設定をすべてのアクティブインスタンスグループに適用する)] を選択します。
- 変更を保存します。

## CLI を使用したインスタンスグループの再設定

`modify-instance-groups` コマンドを使用して、実行中のクラスター内のインスタンスグループに新しい設定を指定します。

### Note

次の例では、`<j-2AL4XXXXXX5T9>` をクラスター ID に、`<ig-1xxxxxxxx9>` をインスタンスグループ ID に置き換えてください。

### Example - インスタンスグループの設定を置き換える

次の例では、という設定 JSON ファイルを参照して、インスタンスグループの YARN NodeManager ディスクヘルスチェッカーのプロパティ `instanceGroups.json` を編集します。

- 設定分類を準備し、コマンドを実行するのと同じディレクトリに `instanceGroups.json` として保存します。

```
[
  {
    "InstanceGroupId": "<ig-1xxxxxxxx9>",
    "Configurations": [
      {
        "Classification": "yarn-site",
        "Properties": {
          "yarn.nodemanager.disk-health-checker.enable": "true",
```

```

        "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
      },
      "Configurations":[]
    }
  ]
}
]

```

2. 以下のコマンドを実行します。

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

### Example - インスタンスグループに設定を追加する

インスタンスグループに設定を追加する場合は、そのインスタンスグループに対して前に指定した設定すべてを新しい ModifyInstanceGroup リクエストに含める必要があります。含めない場合、前に指定した設定が削除されます。

次の例では、YARN NodeManager 仮想メモリチェッカーのプロパティを追加します。設定には、YARN NodeManager ディスクヘルスチェッカーに以前に指定した値も含まれるため、値は上書きされません。

1. instanceGroups.json に次の内容を準備し、コマンドを実行するのと同じディレクトリに保存します。

```

[
  {
    "InstanceId":"<ig-1xxxxxxx9>",
    "Configurations":[
      {
        "Classification":"yarn-site",
        "Properties":{
          "yarn.nodemanager.disk-health-checker.enable":"true",
          "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0",
          "yarn.nodemanager.vmem-check-enabled":"true",
          "yarn.nodemanager.vmem-pmem-ratio":"3.0"
        }
      }
    ]
  }
]

```

```

    },
    "Configurations":[]
  }
]
}
]

```

2. 以下のコマンドを実行します。

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

### Example — マージタイプの再設定を使用してインスタンスグループに設定を追加する

デフォルトの上書き再設定によってインスタンスグループに設定を追加する場合は、そのインスタンスグループに以前指定したすべての設定を、新しい ModifyInstanceGroup リクエストに含める必要があります。そうしない場合、上書きによって、以前指定した設定が削除されます。マージの再設定では、この操作は不要です。ただし、リクエストでは、新しい設定のみ指定する必要があります。

次の例では、YARN NodeManager 仮想メモリチェッカーのプロパティを追加します。これはマージタイプの再設定であるため、YARN NodeManager ディスクヘルスチェッカーに以前に指定した値は上書きされません。

1. instanceGroups.json に次の内容を準備し、コマンドを実行するのと同じディレクトリに保存します。

```

[
  {
    "InstanceId": "<ig-1xxxxxxx9>",
    "ReconfigurationType": "MERGE",
    "Configurations": [
      {
        "Classification": "yarn-site",
        "Properties": {
          "yarn.nodemanager.vmem-check-enabled": "true",
          "yarn.nodemanager.vmem-pmem-ratio": "3.0"
        }
      },
      {
        "Classification": "yarn-site",
        "Properties": {
          "yarn.nodemanager.vmem-check-enabled": "true",
          "yarn.nodemanager.vmem-pmem-ratio": "3.0"
        }
      }
    ]
  }
]

```

```
]
```

2. 以下のコマンドを実行します。

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \  
--instance-groups file://instanceGroups.json
```

### Example - インスタンスグループの設定を削除する

インスタンスグループの設定を削除するには、前の設定を除外する新しい再設定リクエストを送信します。

#### Note

最初のクラスター設定に対して実行できるのは上書きのみです。この設定を削除することはできません。

例えば、前の例から YARN NodeManager ディスクヘルスチェッカーの設定を削除するには、次の内容 `instanceGroups.json` の新しい を送信します。

```
[  
  {  
    "InstanceId": "<ig-1xxxxxxx9>",  
    "Configurations": [  
      {  
        "Classification": "yarn-site",  
        "Properties": {  
          "yarn.nodemanager.vmem-check-enabled": "true",  
          "yarn.nodemanager.vmem-pmem-ratio": "3.0"  
        },  
        "Configurations": []  
      },  
      {}  
    ]  
  }  
]
```



**Note**

最後の再設定リクエストに含まれる設定をすべて削除するには、空の設定の配列を指定して再設定リクエストを送信します。例:

```
[
  {
    "InstanceGroupId": "<ig-1xxxxxxx9>",
    "Configurations": []
  }
]
```

**Example — 1 つのリクエストでインスタンスグループの再設定とサイズ変更を実行する**

同じリクエストでインスタンスグループの再設定とサイズ変更を行う方法を次の JSON 例に示します。

```
[
  {
    "InstanceGroupId": "<ig-1xxxxxxx9>",
    "InstanceCount": 5,
    "EC2InstanceIdsToTerminate": ["i-123"],
    "ForceShutdown": true,
    "ShrinkPolicy": {
      "DecommissionTimeout": 10,
      "InstanceResizePolicy": {
        "InstancesToTerminate": ["i-123"],
        "InstancesToProtect": ["i-345"],
        "InstanceTerminationTimeout": 20
      }
    },
    "Configurations": [
      {
        "Classification": "yarn-site",
        "Configurations": [],
        "Properties": {
          "yarn.nodemanager.disk-health-checker.enable": "true",
          "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-percentage": "100.0"
        }
      }
    ]
  }
]
```

```
    }  
  ]  
}  
]
```

## Java SDK を使用したインスタンスグループの再設定

### Note

次の例では、`<j-2AL4XXXXXX5T9>` をクラスター ID に、`<ig-1xxxxxxx9>` をインスタンスグループ ID に置き換えてください。

次のコードスニペットでは、AWS SDK for Javaを使用して、インスタンスグループに対する新しい設定を指定します。

```
AWSCredentials credentials = new BasicAWSCredentials("access-key", "secret-key");  
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);  
  
Map<String,String> hiveProperties = new HashMap<String,String>();  
hiveProperties.put("hive.join.emit.interval","1000");  
hiveProperties.put("hive.merge.mapfiles","true");  
  
Configuration configuration = new Configuration()  
    .withClassification("hive-site")  
    .withProperties(hiveProperties);  
  
InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()  
    .withInstanceId("<ig-1xxxxxxx9>")  
    .withReconfigurationType("MERGE");  
    .withConfigurations(configuration);  
  
ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()  
    .withClusterId("<j-2AL4XXXXXX5T9>")  
    .withInstanceGroups(igConfig);  
  
emr.modifyInstanceGroups(migRequest);
```

次のコードスニペットでは、空の設定の配列を指定して、インスタンスグループに対して前に指定した設定を削除します。

```
List<Configuration> configurations = new ArrayList<Configuration>();

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
    .withInstanceId("<ig-1xxxxxxx9>")
    .withConfigurations(configurations);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

## インスタンスグループの再設定のトラブルシューティング

インスタンスグループの再設定プロセスが失敗した場合、Amazon EMR は再設定を元に戻して、Amazon CloudWatch イベントを使用して失敗メッセージをログに記録します。イベントには、再設定失敗の簡単な概要が記載されます。ここに、再構成が失敗したインスタンスと、対応する失敗メッセージがリストされます。失敗メッセージの例を次に示します。

```
The reconfiguration operation for instance group ig-1xxxxxxx9 in Amazon EMR
cluster j-2AL4XXXXXX5T9 (ExampleClusterName)
failed at 2021-01-01 00:00 UTC and took 2 minutes to fail. Failed configuration version
is example12345.
Failure message: Instance i-xxxxxxx1, i-xxxxxxx2, i-xxxxxxx3 failed with message "This
is an example failure message".
```

再設定失敗に関する詳細データを収集するには、ノードのプロビジョニングログを確認します。これは、次のようなメッセージを受信したときに特に有用です。

```
i-xxxxxxx1 failed with message "Unable to complete transaction and some changes were
applied."
```

## On the node

ノードに接続してノードのプロビジョニングログにアクセスする

1. SSH を使用して、再設定が失敗したノードに接続します。手順については、Amazon EC2 Linux インスタンス用ユーザーガイドの [Linux インスタンスへの接続](#) を参照してください。
2. ノードのプロビジョニングログファイルが含まれる次のディレクトリに移動します。

```
/mnt/var/log/provision-node/
```

3. reports サブディレクトリを開き、再設定のノードのプロビジョニングレポートを検索します。reports ディレクトリでは、再設定のバージョン番号、Universally Unique Identifier (UUID)、Amazon EC2 インスタンスの IP アドレス、およびタイムスタンプによってログが整理されます。各レポートは、再設定プロセスに関する詳細情報が含まれる圧縮された YAML ファイルです。

レポートファイルの名前とパスの例を示します。

```
/reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxxxx/ip-10-73-xxx-xxx.ec2.internal/202104061715.yaml.gz
```

4. レポートは、次の例に示すように、zless などのファイルビューワーを使用して確認できます。

```
zless 202104061715.yaml.gz
```

## Amazon S3

Amazon S3 を使用してノードのプロビジョニングログにアクセスするには

1. にサインイン AWS Management Console し、<https://console.aws.amazon.com/s3/> で Amazon S3 コンソールを開きます。
2. ログファイルをアーカイブするようにクラスターを設定したときに指定した Amazon S3 バケットを開きます。
3. ノードのプロビジョニングログファイルが含まれる次のフォルダに移動します。

```
DOC-EXAMPLE-BUCKET/elasticmapreduce/<cluster id>/node/<instance id>/provision-node/
```

- reports フォルダを開き、再設定のノードのプロビジョニングレポートを検索します。reports フォルダでは、再設定のバージョン番号、Universally Unique Identifier (UUID)、Amazon EC2 インスタンスの IP アドレス、およびタイムスタンプによってログが整理されます。各レポートは、再設定プロセスに関する詳細情報が含まれる圧縮された YAML ファイルです。

レポートファイルの名前とパスの例を次に示します。

```
/reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxxxx/ip-10-73-xxx-xxx.ec2.internal/202104061715.yaml.gz
```

- ログファイルを表示するには、Amazon S3 からローカルマシンにログファイルをテキストファイルとしてダウンロードします。手順については、「[オブジェクトのダウンロード](#)」を参照してください。

各ログファイルには、関連する再設定に関する詳細なプロビジョニングレポートが含まれています。エラーメッセージ情報を検索するには、レポートの err ログレベルを検索します。レポートの形式は、クラスターの Amazon EMR のバージョンによって異なります。

5.32.0 および 6.2.0 より前の Amazon EMR リリースバージョンのエラー情報を次の例に示します。

```
- !ruby/object:Puppet::Util::Log
  level: !ruby/sym err
  tags:
    - err
  message: "Example detailed error message."
  source: Puppet
  time: 2021-01-01 00:00:00.000000 +00:00
```

Amazon EMR リリースバージョン 5.32.0 および 6.2.0 以降では、代わりに次の形式が使用されます。

```
- level: err
  message: 'Example detailed error message.'
  source: Puppet
  tags:
    - err
  time: '2021-01-01 00:00:00.000000 +00:00'
  file:
  line:
```

## 機密設定データを AWS Secrets Manager に保存する

Amazon EMR では、カスタム設定データ (DescribeCluster や ListInstanceGroups など) を発行する API オペレーションをプレーンテキストで記述して一覧表示できます。Amazon EMR はと統合 AWS Secrets Manager されているため、データを Secrets Manager に保存し、設定でシークレット ARN を使用できます。この方法を取ると、機密性の高い設定データをプレーンテキストで Amazon EMR に渡して外部 API に公開しなくても済みます。キーと値のペアに Secrets Manager に保存されているシークレットの ARN が含まれていることを示すと、Amazon EMR は設定データをクラスターに送信するときにこのシークレットを取得し、そのシークレット ARN が取得されます。Amazon EMR は、外部 API を使用して設定を表示する場合、そのアノテーションを送信されません。

### シークレットを作成する

シークレットを作成するには、「AWS Secrets Manager ユーザーガイド」の「[AWS Secrets Manager シークレットを作成する](#)」の手順に従います。ステップ 3 では、[プレーンテキスト] フィールドを選択して機密値を入力する必要があります。

Secrets Manager では、最大 65536 バイトをシークレットで指定できますが、Amazon EMR では、プロパティキー (アノテーションを除く) と取得するシークレット値を合わせた長さが 1024 文字に制限されていることに注意してください。

### Amazon EMR にアクセス権を付与してシークレットを取得する

Amazon EMR では、IAM サービスロールを使用してクラスターをプロビジョニングおよび管理します。Amazon EMR に許可するアクションは Amazon EMR のサービスロールによって定義します。これにより、リソースをプロビジョニングすると共に、クラスター内で稼働している Amazon EC2 インスタンスのコンテキスト上、実行されないサービスレベルのタスクを実行できます。サービスロールの詳細については、「[Amazon EMR のサービスロール \(EMR ロール\)](#)」と「[IAM ロールのカスタマイズ](#)」を参照してください。

Amazon EMR が Secrets Manager からシークレット値を取得できるようにするには、クラスターを起動するときに Amazon EMR ロールに次のポリシーステートメントを追加します。

```
{
  "Sid": "AllowSecretsRetrieval",
  "Effect": "Allow",
  "Action": "secretsmanager:GetSecretValue",
  "Resource": [
```

```
    "arn:aws:secretsmanager:<region>:<aws-account-id>:secret:<secret-name>"
  ]
}
```

カスタマー管理の でシークレットを作成する場合は AWS KMS key、使用するキーの Amazon EMR ロールにも `kms:Decrypt` アクセス許可を追加する必要があります。詳細については、AWS Secrets Manager ユーザーガイドの「[AWS Secrets Manager に対する認証とアクセスコントロール](#)」を参照してください。

## シークレットを設定分類で使用する

EMR.secret@ アノテーションを任意の設定プロパティに追加して、そのキーと値のペアに、Secrets Manager に保存しているシークレットの ARN が含まれていることを示すことができます。

次の例は、設定分類でシークレット ARN を指定する方法を示しています。

```
{
  "Classification": "core-site",
  "Properties": {
    "presto.s3.access-key": "<sensitive-access-key>",
    "EMR.secret@presto.s3.secret-key": "arn:aws:secretsmanager:<region>:<aws-account-id>:secret:<secret-name>"
  }
}
```

クラスターの作成時にアノテーション付きの設定を送信すると、Amazon EMR によって設定プロパティが検証されます。設定が有効な場合、Amazon EMR は、設定からアノテーションを削除し、Secrets Manager からシークレットを取得します。これによって、実際の設定を作成し、クラスターに適用します。

```
{
  "Classification": "core-site",
  "Properties": {
    "presto.s3.access-key": "<sensitive-access-key>",
    "presto.s3.secret-key": "<my-secret-key-retrieved-from-Secrets-Manager>"
  }
}
```

DescribeCluster などのアクションを呼び出すと、Amazon EMR から、クラスターで現在稼働しているアプリケーションの設定が返ります。アプリケーション設定プロパティにシークレット ARN

が含まれるとマークしている場合、DescribeCluster 呼び出しによって返るアプリケーション設定には、シークレット値ではなく ARN が含まれています。これにより、クラスター上でのみシークレット値を目視できるようになります。

```
{
  "Classification":"core-site",
  "Properties":{
    "presto.s3.access-key":"<sensitive-access-key>",
    "presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-id>:secret:<secret-name>"
  }
}
```

## シークレット値を更新する

Amazon EMR は、アタッチされたインスタンスグループが起動、再設定、またはサイズ変更されるたびに、アノテーション付きの設定からシークレット値を取得します。稼働中のクラスターの設定で使用するシークレット値は、Secrets Manager を使用すると変更できます。その場合、更新済みの値を受け取りたい各インスタンスグループに、再設定リクエストを送信できます。インスタンスグループを再設定する方法と、その際に考慮すべき点の詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

## 特定の Java 仮想マシンを使用するようにアプリケーションを設定

Amazon EMR は、リリースごとに、デフォルトの Java 仮想マシン (JVM) のバージョンが異なります。このページでは、さまざまなリリースやアプリケーションに対する JVM サポートについて説明します。

### 考慮事項

アプリケーションでサポートされている Java バージョンについては、[Amazon EMR リリースガイド](#)のアプリケーションのページを参照してください。

- Amazon EMR のクラスター内で実行可能なランタイムバージョンは 1 つのみであり、同じクラスターで、異なるランタイムバージョンを実行し、異なるノードまたはアプリケーションを稼働させることはできません。
- Amazon EMR 7.x の場合、Apache Livy を除く、Java 17 をサポートするアプリケーションのデフォルトの Java 仮想マシン (JVM) は Java 17 です。アプリケーションでサポートされている



JDK バージョンの詳細については、Amazon EMR リリースガイドの対応するリリースページを参照してください。

- Amazon EMR 7.1.0 以降、Flink は をサポートし、デフォルトで Java 17 に設定されています。別のバージョンの Java ランタイムを使用するには、 の設定を上書きします `flink-conf`。Java 8 または Java 11 を使用するように Flink を設定する方法の詳細については、「[Java 11 で実行するように Flink を設定する](#)」を参照してください。
- Amazon EMR 5.x および 6.x の場合、デフォルトの Java 仮想マシン (JVM) は Java 8 です。
  - Amazon EMR リリース 6.12.0 以降の場合、一部のアプリケーションも Java 11 と 17 をサポートしています。
  - Amazon EMR リリース 6.9.0 以降の場合、Trino はデフォルトで Java 17 をサポートしています。Java 17 と Trino の詳細については、Trino のブログ、「[Trino updates to Java 17](#)」を参照してください。

ランタイムバージョンを選択する際には、次に示すアプリケーション固有の考慮事項に留意してください。

#### Java 設定に関する、アプリケーション固有の注意点

アプリケーション	Java の設定に関する注意点
Spark	<p>Spark をデフォルト以外の Java バージョンで稼働させるには、Spark と Hadoop の両方を設定する必要があります。例については、「<a href="#">JVM をオーバーライドする</a>」を参照してください。</p> <ul style="list-style-type: none"> <li>• プライマリインスタンスプロセスの Java ランタイムを更新するには、<code>spark-env</code> の <code>JAVA_HOME</code> を設定します。例えば、<code>spark-submit</code>、<code>spark-shell</code>、Spark History Server などです。</li> <li>• Hadoop の設定を変更して Spark エグゼキューターと YARN ApplicationMaster の Java ランタイムを更新してください。</li> </ul>

アプリケーション	Java の設定に関する注意点
Spark RAPIDS	Spark 用に設定されている Java バージョンで RAPIDS を実行できます。
Iceberg	Iceberg は、それを使用するアプリケーションの設定済み Java バージョンで稼働させることができます。
差分	Delta は、それを使用するアプリケーションの設定済み Java バージョンで稼働させることができます。
Hudi	Hudi は、それを使用するアプリケーションの設定済み Java バージョンで稼働させることができます。
Hadoop	Hadoop の JVM を更新するには、 <code>hadoop-env</code> を変更します。例については、「 <a href="#">JVM をオーバーライドする</a> 」を参照してください。
[Hive]	Hive の Java バージョンを 11 または 17 に設定するには、Hadoop の JVM 設定で、使用する Java バージョンを指定します。
HBase	HBase の JVM を更新するには、 <code>hbase-env</code> を変更します。デフォルトの場合、Amazon EMR では、Hadoop の JVM 設定に基づいて HBase の JVM が設定されます。ただし、 <code>hbase-env</code> の設定をオーバーライドした場合を除きます。例については、「 <a href="#">JVM をオーバーライドする</a> 」を参照してください。

アプリケーション	Java の設定に関する注意点
Flink	Flink の JVM を更新するには、 <code>flink-conf</code> を変更します。デフォルトの場合、Amazon EMR では、Hadoop の JVM 設定に基づいて Flink の JVM が設定されます。ただし、 <code>flink-conf</code> の設定をオーバーライドした場合を除きます。詳細については、「 <a href="#">Flink が Java 11 で実行されるよう設定する</a> 」を参照してください。
Oozie	Oozie が Java 11 または 17 で稼働するように構成するには、Oozie サーバーと Oozie LauncherAM ランチャー AM の設定を行い、クライアント側の実行ファイルとジョブ設定を変更します。 <code>EmbeddedOozieServer</code> を設定して、Java 17 で稼働させることもできます。詳細については、「 <a href="#">Oozie 用の Java バージョンの設定</a> 」を参照してください。
Pig	Pig は Java 8 のみをサポートしています。Hadoop で Java 11 または 17 を使用し、それと同じクラスターで Pig を稼働させることはできません。

## JVM をオーバーライドする

Amazon EMR リリースの JVM 設定をオーバーライドするには (例えば、Amazon EMR リリース 6.12.0 を使用するクラスターで Java 17 を使用するなど)、`JAVA_HOME` をその環境分類に設定します。この場合の環境分類は `application-env` であり、Flink を除くすべてのアプリケーションに適用されます。Flink の環境分類は `flink-conf` です。Flink の Java ランタイムを設定する手順については、「[Flink が Java 11 で実行されるよう設定する](#)」を参照してください。

### トピック

- [JVM 設定をオーバーライドし Apache Spark を使用する](#)
- [JVM 設定をオーバーライドし Apache HBase を使用する](#)

- [JVM 設定をオーバーライドし Apache Hadoop と Hive を使用する](#)

## JVM 設定をオーバーライドし Apache Spark を使用する

Amazon EMR リリース 6.12 以降で Spark を使用し、クラスターモードで送信するドライバーを作成すると、そのドライバーは Java 8 を使用します。ただし、エグゼキューターが Java 11 または 17 を使用するよう環境を設定できます。Amazon EMR リリース 5.x より前のリリースで Spark を使用し、クラスターモードで送信するドライバーを作成すると、ドライバーは Java 7 を使用しますが、エグゼキューターが Java 8 を使用するよう環境を設定できます。

Spark の JVM をオーバーライドするには、Hadoop 分類と Spark 分類の両方を設定すると良いでしょう。

```
{
  "Classification": "hadoop-env",
    "Configurations": [
      {
        "Classification": "export",
          "Configurations": [],
          "Properties": {
            "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
          }
        },
      ],
    "Properties": {}
  },
  {
    "Classification": "spark-env",
      "Configurations": [
        {
          "Classification": "export",
            "Configurations": [],
            "Properties": {
              "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
            }
          },
        ],
      "Properties": {}
    }
  }
```

## JVM 設定をオーバーライドし Apache HBase を使用する

HBase で Java 11 が使用されるように設定するには、クラスターの起動時に次の設定を行います。

```
[
  {
    "Classification": "hbase-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/jre-11"
        }
      }
    ],
    "Properties": {}
  }
]
```

## JVM 設定をオーバーライドし Apache Hadoop と Hive を使用する

次の例は、JVM をバージョン 17 に設定し、Hadoop と Hive を利用可能にする方法を示しています。

```
[
  {
    "Classification": "hadoop-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/jre-17"
        }
      }
    ],
    "Properties": {}
  }
]
```

## サービスポート

以下は YARN および HDFS サービスポートです。これらの設定には、Hadoop のデフォルトが反映されます。特に文言化されていない限り、他のアプリケーションサービスはデフォルトのポートでホストされます。詳細については、アプリケーションのプロジェクトドキュメントを参照してください。

### YARN と HDFS のポート設定

設定	ホスト名 / ポート
<code>fs.default.name</code>	デフォルト ( <code>hdfs://<i>emrDeterminedIP</i>:8020</code> )
<code>dfs.datanode.address</code>	デフォルト ( <code>0.0.0.0:50010</code> )
<code>dfs.datanode.http.address</code>	デフォルト ( <code>0.0.0.0:50075</code> )
<code>dfs.datanode.https.address</code>	デフォルト ( <code>0.0.0.0:50475</code> )
<code>dfs.datanode.ipc.address</code>	デフォルト ( <code>0.0.0.0:50020</code> )
<code>dfs.http.address</code>	デフォルト ( <code>0.0.0.0:50070</code> )
<code>dfs.https.address</code>	デフォルト ( <code>0.0.0.0:50470</code> )
<code>dfs.secondary.http.address</code>	デフォルト ( <code>0.0.0.0:50090</code> )
<code>yarn.nodemanager.address</code>	デフォルト ( <code>\${yarn.nodemanager.hostname}:0</code> )
<code>yarn.nodemanager.localizer.address</code>	デフォルト ( <code>\${yarn.nodemanager.hostname}:8040</code> )
<code>yarn.nodemanager.webapp.address</code>	デフォルト ( <code>\${yarn.nodemanager.hostname}:8042</code> )
<code>yarn.resourcemanager.address</code>	デフォルト ( <code>\${yarn.resourcemanager.hostname}:8032</code> )

設定	ホスト名 / ポート
<code>yarn.resourcemanager.admin.address</code>	デフォルト ( <code>\${yarn.resourcemanager.hostname}:8033</code> )
<code>yarn.resourcemanager.resource-tracker.address</code>	デフォルト ( <code>\${yarn.resourcemanager.hostname}:8031</code> )
<code>yarn.resourcemanager.scheduler.address</code>	デフォルト ( <code>\${yarn.resourcemanager.hostname}:8030</code> )
<code>yarn.resourcemanager.webapp.address</code>	デフォルト ( <code>\${yarn.resourcemanager.hostname}:8088</code> )
<code>yarn.web-proxy.address</code>	デフォルト (no-value )
<code>yarn.resourcemanager.hostname</code>	<i>emrDeterminedIP</i>

### Note

*emrDeterminedIP* とは、Amazon EMR コントロールプレーンで生成された IP アドレスのことです。新しいバージョンでは、この規則は削除されています ( `yarn.resourcemanager.hostname` および `fs.default.name` 設定は除く)。

## アプリケーションユーザー

アプリケーションでは、プロセスを独自のユーザーとして実行します。例えば、Hive JVMsユーザーとして実行mapredされhive、MapReduce JVMsとして実行されます。これを、次のプロセスステータスの例に示します。

```

USER      PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
hive      6452  0.2  0.7 853684 218520 ?        S1   16:32   0:13 /usr/lib/jvm/
java-openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-
metastore.log -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop
hive      6557  0.2  0.6 849508 202396 ?        S1   16:32   0:09 /usr/lib/jvm/java-
openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-server2.log
-Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop/1

```

```
hbase      6716  0.1  1.0 1755516 336600 ?      S1   Jun21   2:20 /usr/lib/jvm/java-  
openjdk/bin/java -Dproc_master -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:  
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/  
hbase      6871  0.0  0.7 1672196 237648 ?      S1   Jun21   0:46 /usr/lib/jvm/java-  
openjdk/bin/java -Dproc_thrift -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:  
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/  
hdfs       7491  0.4  1.0 1719476 309820 ?      S1   16:32   0:22 /usr/lib/jvm/java-  
openjdk/bin/java -Dproc_namenode -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-hdfs -  
Dhadoop.log.file=hadoop-hdfs-namenode-ip-10-71-203-213.log -Dhadoo  
yarn       8524  0.1  0.6 1626164 211300 ?      S1   16:33   0:05 /usr/lib/jvm/java-  
openjdk/bin/java -Dproc_proxyserver -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn -  
Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-yarn-  
yarn       8646  1.0  1.2 1876916 385308 ?      S1   16:33   0:46 /usr/lib/jvm/java-  
openjdk/bin/java -Dproc_resourceanager -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn  
-Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-y  
mapred     9265  0.2  0.8 1666628 260484 ?      S1   16:33   0:12 /usr/lib/jvm/java-  
openjdk/bin/java -Dproc_historyserver -Xmx1000m -Dhadoop.log.dir=/usr/lib/hadoop/logs -  
Dhadoop.log.file=hadoop.log -Dhadoop.home.dir=/usr/lib/hadoop
```



# Amazon EMR アーティファクトリポジトリを使用して依存関係を確認する

Amazon EMR リリースバージョン 5.18.0 以降では、Amazon EMR アーティファクトリポジトリを使用し、特定の Amazon EMR リリースバージョンで利用可能なライブラリおよび依存関係の正確なバージョンに対して Apache Hive および Apache Hadoop のジョブコードをビルドできます。リポジトリの Amazon EMR アーティファクトに対してビルドすると、ジョブがビルドされるライブラリのバージョンがクラスターでの実行時に提供されるバージョンと正確に同じになるため、ランタイムクラッシュの問題を回避できます。現在、Amazon EMR アーティファクトは Maven ビルドでのみ使用可能です。

アーティファクトリポジトリにアクセスするには、リポジトリ URL を Maven 設定ファイルまたは特定のプロジェクトの `pom.xml` 設定ファイルに追加します。その後、プロジェクト設定に依存関係を指定できます。依存関係のバージョンについては、[Amazon EMR 5.x リリースバージョン](#) にある目的のリリースのコンポーネントバージョンの下にリストされているバージョンを使用します。例えば、最新の Amazon EMR リリースのコンポーネントバージョンは [the section called “5.36.2 コンポーネントバージョン”](#) にあります。プロジェクトのアーティファクトがコンポーネントバージョンにリストされていない場合は、そのリリースの Hive と Hadoop にリストされているバージョンを指定します。例えば、Amazon EMR リリースバージョン 5.18.0 の Hadoop コンポーネントの場合、バージョンは `2.8.4-amzn-1` です。

アーティファクトリポジトリ URL の構文は次のとおりです。

```
https://s3-endpoint/region-ID-emr-artifacts/emr-release-label/repos/maven/
```

- **s3-endpoint** はリポジトリのリージョンの Amazon Simple Storage Service (Amazon S3) エンドポイントで、**region-ID** は対応するリージョンです。例えば、`s3.us-west-1.amazonaws.com` と `us-west-1` です。詳細については、「Amazon Web Services 全般のリファレンス」の Amazon S3 エンドポイントを参照してください。リージョン間でアーティファクトに違いはないため、開発環境に対して最も便利なリージョンを指定することができます。
- **emr-release-label** は、コードを実行する Amazon EMR クラスターのリリースラベルです。リリースラベルは `emr-x.x.x` 形式 (`emr-5.36.2` など) です。EMR リリースシリーズには、複数のリリースが含まれる場合があります。たとえば、EMR リリースバージョン 5.24.1 を使用している場合、アーティファクトリポジトリ URL では、5.24 シリーズの初回 EMR リリースラベル `emr-5.24.0` を使用します。

```
https://s3-endpoint/region-ID-emr-artifacts/emr-5.24.0/repos/maven/
```

## Example Maven pom.xml の設定

次の pom.xml の例では、us-west-1 のアーティファクトリポジトリを使用して、emr-5.18.0 Apache Hadoop および Apache Hive アーティファクトに対してビルドするように Maven プロジェクトを設定します。スナップショットのバージョンはアーティファクトリポジトリでは使用できないため、スナップショットは pom.xml で無効になります。次の例の省略記号 (...) は、他の設定パラメータの省略を示します。Maven プロジェクトにこれらをコピーしないでください。

```
<project>
  ...
  <repositories>
    ...
    <repository>
      <id>emr-5.18.0-artifacts</id>
      <name>EMR 5.18.0 Releases Repository</name>
      <releases>
        <enabled>>true</enabled>
      </releases>
      <snapshots>
        <enabled>>false</enabled>
      </snapshots>
      <url>https://s3.us-west-1.amazonaws.com/us-west-1-emr-artifacts/emr-5.18.0/repos/
maven/</url>
    </repository>
    ...
  </repositories>
  ...
  <dependencies>
    ...
    <dependency>
      <groupId>org.apache.hive</groupId>
      <artifactId>hive-exec</artifactId>
      <version>2.3.3-amzn-2</version>
    </dependency>
    <dependency>
      <groupId>org.apache.hadoop</groupId>
      <artifactId>hadoop-common</artifactId>
      <version>2.8.4-amzn-1</version>
    </dependency>
  </dependencies>
</project>
```

```
...  
</dependencies>  
  
</project>
```

## EMR ファイルシステム (EMRFS)

EMR ファイルシステム (EMRFS) は、すべての Amazon EMR クラスターが Amazon EMR と Amazon S3 の間で通常のファイルを直接読み書きするために使用する HDFS の実装です。EMRFS には Hadoop で使用するために Amazon S3 に永続的なデータを保存するという利便性がある一方で、データの暗号化などの機能も用意されています。

データの暗号化により、EMRFS が Amazon S3 に書き込むオブジェクトが暗号化され、Amazon S3 の暗号化されたオブジェクトを EMRFS で操作できるようになります。Amazon EMR リリース 4.8.0 以降のバージョンを使用している場合は、セキュリティ設定を使用して、Amazon S3 の EMRFS オブジェクトの暗号化を、他の暗号化設定と合わせてセットアップできます。詳細については「[暗号化オプション](#)」を参照してください。これまでのリリースバージョンの Amazon EMR を使用している場合は、手動で暗号化を設定できます。詳細については、「[EMRFS プロパティを使用して Amazon S3 の暗号化を指定する](#)」を参照してください。

Amazon S3 には、すべての AWS リージョンでの GET、PUT、および LIST の全操作について、書き込み後読み取りの強力な整合性維持機能があります。つまり、EMRFS を使用して書き込んだ内容が Amazon S3 から読み取る内容であり、パフォーマンスにも影響はありません。詳細については、「[Amazon S3 のデータ整合性モデル](#)」を参照してください。

Amazon EMR リリース 5.10.0 以降のバージョンを使用するときは、Amazon S3 のクラスターユーザー、グループ、または EMRFS データの場所に基づいて、Amazon S3 への EMRFS リクエストに異なる IAM ロールを使用できます。詳細については、「[Amazon S3 への EMRFS リクエストの IAM ロールを設定する](#)」を参照してください。

### Warning

Apache Spark ジョブを実行する Amazon EMR クラスターの投機的実行を有効にする前に、次の情報を確認してください。

EMRFS には、EMRFS S3-optimized コミッターが含まれています。これは、EMRFS を使用するとき Amazon S3 にファイルを書き込むために最適化された OutputCommitter 実装です。Amazon S3 にデータを書き込み、EMRFS S3 最適化コミッターを使用しないアプリケーションで Apache Spark 投機的実行機能をオンにすると、「[SPARK-10063](#)」で説明されているデータの正確性の問題が発生する可能性があります。これは、Amazon EMR リリース 5.19 より前の Amazon EMR バージョンを使用しているか、ORC や CSV などのファイル形式で Amazon S3 に書き込んでいる場合に発生する可能性があります。EMRFS S3 最適化コミッターは、これらの形式をサポートしていません。EMRFS S3 最適化コミッターを使用す

るための要件の詳細なリストについては、「[EMRFS S3 最適化コミッターの要件](#)」を参照してください。

EMRFS 直接書き込みは、通常、次を書き込むときなどの、EMRFS S3 最適化コミッターがサポートされていないときに使用されます。

- ORC やテキストなどの、Parquet 以外の出力形式。
- Spark RDD API を使用する Hadoop ファイル。
- Hive を使用する Parquet SerDe。 「[Hive metastore Parquet table conversion](#)」を参照してください。

EMRFS 直接書き込みは、次のシナリオでは使用されません。

- EMRFS S3 最適化コミッターが有効なとき。「[EMRFS S3 最適化コミッターの要件](#)」を参照してください。
- を dynamic partitionOverwriteMode に設定して動的パーティションを記述する場合。
- Hive のデフォルトのパーティションの場所規則に準拠しない場所などの、カスタムのパーティションの場所へ書き込むとき。
- HDFS への書き込みや S3A ファイルシステムの使用など、EMRFS 以外のファイルシステムを使用するとき。

アプリケーションが Amazon EMR 5.14.0 以降で直接書き込みを使用しているかどうかを判別するには、Spark INFO ログ記録を有効にします。「Direct Write: ENABLED」というテキストを含むログ行が Spark ドライバーログまたは Spark エグゼキューターコンテナログのいずれかに存在する場合、Spark アプリケーションが直接書き込みを使用して書き込んだことがわかります。

Amazon EMR クラスターでは、投機的実行はデフォルトで OFF になっています。次の条件が両方ともあてはまる場合は、投機的実行をオンにしないことを強くお勧めします。

- データを Amazon S3 に書き込む場合。
- データを Apache Parquet 以外の形式で、または Apache Parquet 形式だが EMRFS S3 最適化コミッターを使用せずに書き込む場合。

Spark 投機的実行をオンにし、EMRFS 直接書き込みを使用して Amazon S3 にデータを書き込むと、断続的なデータ損失が発生する可能性があります。HDFS にデータを書き

込むか、EMRFS S3 最適化コミッターを使用して Parquet 形式でデータを書き込むときは、Amazon EMR が直接書き込みを使用しないため、この問題は発生しません。EMRFS 直接書き込みを使用する形式で Spark から Amazon S3 にデータを書き込み、かつ投機的実行を使用する必要がある場合は、HDFS に書き込み、S3DistCp を使用して出力ファイルを Amazon S3 に転送することをお勧めします。

## トピック

- [整合性のあるビュー](#)
- [Amazon S3 内の EMRFS データへのアクセスを許可する](#)
- [デフォルト AWS Security Token Service エンドポイントの管理](#)
- [EMRFS プロパティを使用して Amazon S3 の暗号化を指定する](#)

## 整合性のあるビュー

### Warning

2023 年 6 月 1 日以降、その後リリースされる Amazon EMR では、EMRFS の整合性のあるビューに標準サポートは提供されません。既存のリリースの場合、EMRFS の整合性のあるビューは、継続して利用できます。

2020 年 12 月 1 日の Amazon S3 の強力な read-after-write 整合性のリリースにより、Amazon EMR クラスターで EMRFS 整合性ビュー (EMRFS CV) を使用する必要がなくなりました。EMRFS CV は、Amazon EMR クラスターが Amazon S3 オブジェクトのリストと read-after-write 整合性をチェックできるようにするオプション機能です。クラスターを作成し、EMRFS CV が有効になっている場合、Amazon EMR は Amazon DynamoDB データベースを作成し、S3 オブジェクトのリストと read-after-write 整合性を追跡するために使用するオブジェクトメタデータを保存します。今回のリリースにより、EMRFS CV をオフにして、この機能が使用する DynamoDB データベースを削除できるようになりました。これで、追加のコストも発生しません。次の手順に従うと、CV 機能の確認と無効化に加え、その機能が使用する DynamoDB データベースの削除を行えます。

EMRFS CV 機能の使用状況を確認するには

1. [設定] タブに移動します。クラスターを以下のように設定している場合、EMRFS CV が使用されます。

```
Classification=emrfs-site,Property=fs.s3.consistent,Value=true
```

2. または、を使用して [describe-cluster API](#) でクラスターを AWS CLI 記述します。出力に `fs.s3.consistent: true` が指定されていると、クラスターは EMRFS CV を使用します。

Amazon EMR クラスターの EMRFS CV を無効にするには

EMRFS CV の機能を無効にするには、次に示す 3 つのオプションのいずれかを使用します。これらのオプションは、テスト環境でテストした後に、実稼働環境に適用する必要があります。

1. 既存のクラスターを停止し、EMRFS CV オプションを使用せずに新規クラスターを開始するには。
  - a. クラスターの停止前に、データを確実にバックアップし、ユーザーに停止を通知してください。
  - b. クラスターを停止するには、「[クラスターを終了する](#)」の手順に従います。
  - c. Amazon EMR コンソールからクラスターを新規作成する場合は、[詳細オプション] に移動します。[ソフトウェア設定の編集] セクションで、EMRFS CV を有効にするオプションの選択を解除します。[EMRFS の整合性のビュー] のチェックボックスが使用可能な場合は、チェックを外したままにします。
  - d. AWS CLI を使用して [create-cluster API で新しいクラスターを作成する場合は](#)、EMRFS CV をオンにする `--emrfs` オプションを使用しないでください。
  - e. SDK またはを使用して新しいクラスター AWS CloudFormation を作成する場合は、「[整合性のあるビューを設定する](#)」に記載されている設定を使用しないでください。
2. クラスターを複製して EMRFS CV を削除するには
  - a. Amazon EMR コンソールで、EMRFS CV を使用するクラスターを選択します。
  - b. [クラスターの詳細] ページの上部で、[クローン] を選択します。
  - c. [前へ] を選択して、[ステップ 1: ソフトウェアとステップ] に移動します。
  - d. [ソフトウェア設定の編集] で EMRFS CV を削除します。[設定を編集する] で、`emrfs-site` 分類内にある次の設定を削除します。S3 バケットから JSON を読み込む場合は、S3 オブジェクトを変更する必要があります。

```
[  
  {"classification":  
    "emrfs-site",
```

```
"properties": {
  "fs.s3.consistent.retryPeriodSeconds": "10",
  "fs.s3.consistent": "true",
  "fs.s3.consistent.retryCount": "5",
  "fs.s3.consistent.metadata.tableName": "EmrFSMetadata"
}
}
]
```

### 3. インスタンスグループを使用するクラスターから EMRFS CV を削除するには

- a. 1つの EMR クラスターが EMRFS CV に関連付けられた DynamoDB テーブルを使用しているかどうか、または複数のクラスターがテーブルを共有しているかどうかを確認するには、次のコマンドを使用します。「[整合性のあるビューを設定する](#)」の説明どおり、テーブル名は `fs.s3.consistent.metadata.tableName` で指定します。EMRFS CV で使用するデフォルトのテーブル名は `EmrFSMetadata` です。

```
aws emr describe-cluster --cluster-id j-XXXXX | grep
fs.s3.consistent.metadata.tableName
```

- b. DynamoDB データベースが別のクラスターと共有されていない場合は、次のコマンドを使用してクラスターを再設定し、EMRFS CV を非アクティブ化します。詳細については、「[実行中のクラスター内のインスタンスグループの再設定](#)」を参照してください。

```
aws emr modify-instance-groups --cli-input-json file://disable-emrfs-1.json
```

このコマンドにより、変更対象のファイルが開きます。次の設定を使用してファイルを変更します。

```
{
  "ClusterId": "j-xxxx",
  "InstanceGroups": [
    {
      "InstanceGroupId": "ig-xxxx",
      "Configurations": [
        {
          "Classification": "emrfs-site",
          "Properties": {
            "fs.s3.consistent": "false"
          }
        },
        "Configurations": []
      ]
    }
  ]
}
```



```
    }  
  ]  
}  
]  
}
```

- c. DynamoDB テーブルが別のクラスターと共有されている場合は、すべてのクラスターの EMRFS CV を一度に無効化します。この操作は、クラスターが S3 の共有クォータにあり、あるオブジェクトを変更しないタイミングで行ってください。

EMRFS CV に関連付けた Amazon DynamoDB リソースを削除するには

Amazon EMR クラスターから EMRFS CV を削除した後、EMRFS CV に関連付けた DynamoDB リソースを削除します。この操作を行わないと、EMRFS CV に関連する DynamoDB 料金が引き続き発生します。

1. DynamoDB テーブルの CloudWatch メトリクスを確認し、テーブルがどのクラスターでも使用されていないことを確認します。
2. DynamoDB テーブルを削除します。

```
aws dynamodb delete-table --table-name <your-table-name>
```

EMRFS CV に関連付けた Amazon SQS リソースを削除するには

1. 不整合の通知が Amazon SQS にプッシュされるようにクラスターを設定すると、SQS キューをすべて削除できます。
2. 「[整合性のあるビューを設定する](#)」の説明どおり `fs.s3.consistent.notification.SQS.queueName` で指定された Amazon SQS キュー名を検索します。デフォルトの場合、キュー名の形式は、EMRFS-Inconsistency-*<j-cluster ID>* です。

```
aws sqs list-queues | grep 'EMRFS-Inconsistency'  
aws sqs delete-queue --queue-url <your-queue-url>
```

## EMRFS CLI の使用を停止するには

- EMRFS CV によって生成されるメタデータは、[EMRFS CLI](#) で管理します。今後リリースされる Amazon EMR では EMRFS CV の標準サポートが終了するため、EMRFS CLI のサポートも終了します。

## トピック

- [整合性のあるビューを有効にする](#)
- [EMRFS の整合性のあるビューが Amazon S3 内のオブジェクトを追跡する方法を理解する](#)
- [再試行ロジック](#)
- [EMRFS の整合性のあるビューのメタデータ](#)
- [CloudWatch および Amazon SQS の整合性通知を設定する](#)
- [整合性のあるビューを設定する](#)
- [EMRFS CLI コマンドリファレンス](#)

## 整合性のあるビューを有効にする

、または設定分類を使用して AWS Management Console、EMRFS の Amazon S3 サーバー側の暗号化または整合性のあるビューを有効にできます。AWS CLI `emrfs-site`

コンソールを使用して整合性のあるビューを設定するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Step 1: Software and Steps] および [Step 2: Hardware] の設定を選択します。
4. [Step 3: General Cluster Settings] の [Additional Options] で、[EMRFS consistent view] を選択します。
5. [EMRFS Metadata store] に、メタデータストアの名前を入力します。デフォルト値は、**EmrFSMetadata**です。EmrFSMetadata テーブルが存在しない場合、自動的に DynamoDB に作成されます。

**Note**

Amazon EMR は、クラスターの終了時に DynamoDB から EMRFS メタデータを自動的に削除しません。

- [Number of retries] に整数値を入力します。不整合が検出された場合、EMRFS はこの回数にわたり、Amazon S3 の呼び出しを試みます。デフォルト値は、**5**です。
- [Retry period (in seconds)] に、整数値を入力します。これは、再試行までに EMRFS が待機する時間です。デフォルト値は、**10**です。

**Note**

それ以降の再試行では、エクスポネンシャルバックオフが使用されます。

を使用して整合性のあるビューを有効にしてクラスターを起動するには AWS CLI

の最新バージョンをインストールすることをお勧めします AWS CLI。最新リリースをダウンロードするには、<https://aws.amazon.com/cli/> を参照してください。

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --instance-type m5.xlarge --instance-count 3 --emrfs  
Consistent=true \  
--release-label emr-7.1.0 --ec2-attributes KeyName=myKey
```

を使用して整合性のあるビューが有効になっているかどうかを確認するには AWS Management Console

- コンソールで整合性のあるビューが有効かどうかを確認するには、[Cluster List] に移動し、クラスター名を選択して [Cluster Details] を表示します。[EMRFS consistent view] フィールドの値は、Enabled または Disabled です。

整合性のあるビューが有効になっているかどうかを確認するには、`emrfs-site.xml` ファイルを調べます。

- 整合性が有効になっているかどうかは、クラスターのマスターノードの `emrfs-site.xml` 設定ファイルを調べることで確認できます。 `fs.s3.consistent` のブール値が `true` に設定されている場合、Amazon S3 が関与するファイルシステム操作に対して整合性のあるビューが有効になっています。

## EMRFS の整合性のあるビューが Amazon S3 内のオブジェクトを追跡する方法を理解する

EMRFS は、オブジェクトに関する情報を EMRFS メタデータに追加することで、Amazon S3 内にそれらのオブジェクトの整合性のあるビューを作成します。EMRFS は、次の場合にこれらの一覧をそのメタデータに追加します。

- Amazon EMR ジョブの実行中に、EMRFS によってオブジェクトが書き込まれた。
- EMRFS CLI を使用してオブジェクトが EMRFS のメタデータと同期化されたか、EMRFS にインポートされた。

EMRFS によって読み取られたオブジェクトはメタデータに自動的に追加されません。EMRFS がオブジェクトを削除されると、リストが EMRFS CLI を使用して消去されるまで、削除された状態でメタデータに残ります。CLI の詳細については、「[EMRFS CLI コマンドリファレンス](#)」を参照してください。EMRFS のメタデータでの一覧消去の詳細については、「[EMRFS の整合性のあるビューのメタデータ](#)」を参照してください。

EMRFS は、すべての Amazon S3 操作について、整合性のあるビューのオブジェクトのセットに関する情報のメタデータを確認します。EMRFS は、これらのいずれかの操作中に Amazon S3 が不整合であることを検出すると、`emrfs-site` 設定プロパティで定義されているパラメータに従って操作を再試行します。EMRFS が再試行した後、`ConsistencyException` を返すか、例外を記録し

てワークフローを継続します。再試行ロジックの詳細については、「[再試行ロジック](#)」を参照してください。たとえば、ログには `ConsistencyExceptions` が見つかります。

- `listStatus`: No Amazon S3 object for metadata item /S3\_bucket/dir/object
- `getFileStatus`: キー `dir/file` はメタデータに存在しますが、Amazon S3 には存在しません

EMRFS の整合性のあるビューで追跡されているオブジェクトを Amazon S3 から直接削除すると、EMRFS はそのオブジェクトを不整合として扱います。これは、そのオブジェクトが Amazon S3 に存在するものとしてメタデータにまだリストされているためです。Amazon S3 で EMRFS が追跡するオブジェクトとメタデータが同期されなくなった場合は、EMRFS CLI の `sync` サブコマンドを使用して、Amazon S3 の状態を反映するようにメタデータをリセットできます。メタデータと Amazon S3 との間の不整合を検出するには、`diff` を使用します。最後に、EMRFS の整合性のあるビューには、メタデータで参照されているオブジェクトのみがあります。同じ Amazon S3 のパスには、追跡されていない他のオブジェクトが存在する可能性があります。EMRFS は、Amazon S3 パスのオブジェクトを一覧表示するときに、メタデータで追跡されているオブジェクトと、その Amazon S3 パス内のオブジェクトのスーパーセットを返します。

## 再試行ロジック

EMRFS は、特定の数の再試行について、メタデータで追跡されるオブジェクトのリストの整合性を検証しようとします。デフォルトは 5 です。再試行の数を超過する場合、`fs.s3.consistent.throwExceptionOnInconsistency` が `false` に設定されていない限り（この場合、整合性がないとして追跡されるオブジェクトのみが記録されます）、失敗を返します。EMRFS は、デフォルトでエクスポネンシャルバックオフ再試行ポリシーを使用しますが、固定ポリシーに設定することもできます。ユーザーは、例外をスローすることなく、残りのジョブに進む前に特定の時間、再試行したい場合があります。その場合は、`fs.s3.consistent.throwExceptionOnInconsistency` を `false` に、`fs.s3.consistent.retryPolicyType` を `fixed` に、`fs.s3.consistent.retryPeriodSeconds` を目的の値に設定することで、これを達成できます。次の例では、整合性を有効にしてクラスターを作成します。これにより、不整合が記録され、10 秒の固定再試行間隔が設定されます。

### Example 再試行期間の一定量への設定

```
aws emr create-cluster --release-label emr-7.1.0 \  
--instance-type m5.xlarge --instance-count 1 \  

```

```
--emrfs Consistent=true,Args=[fs.s3.consistent.throwExceptionOnInconsistency=false,  
fs.s3.consistent.retryPolicyType=fixed,fs.s3.consistent.retryPeriodSeconds=10] --ec2-  
attributes KeyName=myKey
```

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

詳細については、「[整合性のあるビュー](#)」を参照してください。

## IMDS のリージョンの取得呼び出しに対する EMRFS 設定

EMRFS は、IMDS (インスタンスメタデータサービス) を使用して、インスタンスリージョンと Amazon S3、DynamoDB、または AWS KMS エンドポイントを取得します。ただし、IMDS には処理できるリクエストの数に制限があり、その制限を超えるリクエストは失敗します。この IMDS 制限により、EMRFS の初期化が失敗し、クエリまたはコマンドが失敗する可能性があります。emrfs-site.xml で、次のランダム化されたエクスポネンシャルバックオフ再試行メカニズムとフォールバックリージョン設定プロパティを使用して、すべての再試行が失敗するシナリオに対応できます。

```
<property>  
  <name>fs.s3.region.retryCount</name>  
  <value>3</value>  
  <description>  
    Maximum retries that would be attempted to get AWS region.  
  </description>  
</property>  
<property>  
  <name>fs.s3.region.retryPeriodSeconds</name>  
  <value>3</value>  
  <description>  
    Base sleep time in second for each get-region retry.  
  </description>  
</property>  
<property>  
  <name>fs.s3.region.fallback</name>  
  <value>us-east-1</value>  
  <description>
```

```
    Fallback to this region after maximum retries for getting AWS region have been
    reached.
  </description>
</property>
```

## EMRFS の整合性のあるビューのメタデータ

EMRFS の整合性のあるビューは、DynamoDB テーブルを使用して、EMRFS と同期されたか、EMRFS によって作成された Amazon S3 内のオブジェクトの整合性を追跡します。このメタデータはすべての操作（読み取り、書き込み、更新、およびコピー）の追跡に使用され、実際のコンテンツはメタデータには保存されません。このメタデータは、Amazon S3 から受信したオブジェクトやメタデータが、予期しているものと一致するかどうかを検証するために使用されます。この確認により、EMRFS は Amazon S3 に EMRFS が書き込む新しいオブジェクト、read-after-write または EMRFS と同期されたオブジェクトのリストの整合性と整合性をチェックできます。Amazon S3 複数のクラスターで同じメタデータを共有できます。

### メタデータにエントリを追加する方法

メタデータへのエントリの追加には、sync サブコマンドまたは import サブコマンドを使用できます。sync は、パス内の Amazon S3 オブジェクトの状態を反映し、import は厳密に新しいエントリをメタデータに追加するために使用されます。詳細については、「[EMRFS CLI コマンドリファレンス](#)」を参照してください。

### メタデータと Amazon S3 内のオブジェクトとの差異を確認する方法

メタデータと Amazon S3 との差異を確認するには、EMRFS CLI の diff サブコマンドを使用します。詳細については、「[EMRFS CLI コマンドリファレンス](#)」を参照してください。

### メタデータ操作が調整されているかどうかを確認する方法

EMRFS は、その読み取り操作と書き込み操作のメタデータのデフォルトのスループットキャパシティー制限として、それぞれ 500 ユニットと 100 ユニットの単位を設定します。多数のオブジェクトやバケットがあると、操作でこのキャパシティーを超える可能性があります。その時点で DynamoDB は操作を抑制します。たとえば、これらのキャパシティーの上限を超える操作を実行する場合、アプリケーションで EMRFS が ProvisionedThroughputExceededException をスローする場合があります。抑制されると、EMRFS CLI ツールは、操作が終了するか、Amazon EMR から Amazon S3 へのオブジェクト書き込みの最大再試行値に到達するまで、[エクスポネンシャルバックオフ](#)を使用して DynamoDB テーブルへの書き込みを再試行します。

独自のスループットキャパシティー制限を設定できます。ただし、DynamoDB の読み込み操作と書き込み操作には、3000 読み込みキャパシティーユニット (RCU) と 1000 書き込みキャパシティーユ



二ット (WCU) の厳密なパーティション制限があります。抑制による sync 障害を回避するには、読み取り操作のスループットを 3000 RCU 未満に、書き込み操作のスループットを 1000 WCU 未満に制限することをお勧めします。カスタムのスループットキャパシティー制限を設定する手順については、「[整合性のあるビューを設定する](#)」を参照してください。

また、DynamoDB コンソールで EMRFS メタデータの Amazon CloudWatch メトリクスを表示して、スロットリングされた読み取りおよび書き込みリクエストの数を確認することもできます。調整されたリクエストの数がゼロ以外の場合、アプリケーションは、読み取りまたは書き込み操作に割り当てられるスループット容量を増やすことで、メリットが得られる可能性があります。長時間にわたって読み取りと書き込みに割り当てられた最大のスループット容量に操作が近付いていることが表示された場合も、パフォーマンスのメリットが得られる可能性があります。

### 重要な EMRFS 操作のスループット特性

読み取りおよび書き込み操作のデフォルト値は、それぞれ 400 スループット容量ユニット、100 スループット容量ユニットです。以下のパフォーマンス特性は、特定の操作に必要なスループットのヒントを示しています。これらのテストは、単一ノードの m3.large クラスターを使用して実行されました。すべての操作はシングルスレッドです。パフォーマンスは特定のアプリケーション特性によって大きく異なり、ファイルシステム操作を最適化するには、検証が必要になる場合があります。

操作	平均 read-per-second	平均 write-per-second
create ( オブジェクト )	26.79	6.70
delete ( オブジェクト )	10.79	10.79
delete ( 1000 個のオブジェクトを含むディレクトリ )	21.79	338.40
getFileStatus (オブジェクト)	34.70	0
getFileStatus (ディレクトリ)	19.96	0
listStatus ( 1 個のオブジェクトを含むディレクトリ )	43.31	0
listStatus ( 10 個のオブジェクトを含むディレクトリ )	44.34	0



操作	平均 read-per-second	平均 write-per-second
listStatus ( 100 個のオブジェクトを含むディレクトリ )	84.44	0
listStatus (1,000 個のオブジェクトを含むディレクトリ)	308.81	0
listStatus (10,000 個のオブジェクトを含むディレクトリ)	416.05	0
listStatus (100,000 個のオブジェクトを含むディレクトリ)	823.56	0
listStatus ( 1,000,000 個のオブジェクトを含むディレクトリ )	882.36	0
mkdir ( 120 秒間継続 )	24.18	4.03
mkdir	12.59	0
rename ( オブジェクト )	19.53	4.88
rename ( 1000 個のオブジェクトを含むディレクトリ )	23.22	339.34

メタデータストアから古いデータを消去するステップを送信するには

DynamoDB ベースのメタデータの特定エントリを削除することが必要な場合があります。この操作は、テーブルに関連するストレージコストを節約するうえで役立ちます。ユーザーは、EMRFS CLI の delete サブコマンドを使用して、特定のエントリを手動またはプログラムで消去することができます。ただし、メタデータからエントリを削除すると、EMRFS は整合性のチェックを行わなくなります。

ジョブ完了後のプログラムによる消去は、EMRFS CLI でコマンドを実行する最終的なステップをクラスターに送信することにより行います。たとえば、クラスターにステップを送信し、2 日前より古いすべてのエントリを削除するには、次のコマンドを入力します。

```
aws emr add-steps --cluster-id j-2AL4XXXXXX5T9 --steps Name="emrfsCLI",Jar="command-  
runner.jar",Args=["emrfs","delete","--time","2","--time-unit","days"]  
{  
  "StepIds": [  
    "s-B12345678902"  
  ]  
}
```

返された StepId 値を使用して、オペレーションの結果のログを確認します。

## CloudWatch および Amazon SQS の整合性通知を設定する

Amazon S3 の結果整合性の問題に対して、EMRFS で CloudWatch メトリクスと Amazon SQS メッセージを有効にできます。Amazon SQS Amazon S3

### CloudWatch

CloudWatch メトリクスを有効にすると、Amazon S3 の結果整合性のために FileSystem API コールが失敗するたびに、不整合という名前のメトリクスがプッシュされます。

Amazon S3 の結果整合性の問題の CloudWatch メトリクスを表示するには

CloudWatch コンソールで不整合メトリクスを表示するには、EMRFS メトリクスを選択し、JobFlowId/Metric Name ペアを選択します。たとえば、j-162XXXXXXM2CU ListStatusj-162XXXXXXM2CU GetFileStatus などです。

1. <https://console.aws.amazon.com/cloudwatch/> で CloudWatch コンソールを開きます。
2. [Dashboard] の [Metrics] セクションで、[EMRFS] を選択します。
3. ジョブフローメトリクスペインで、1 つ以上の JobFlowId/メトリクス名のペアを選択します。メトリクスのグラフィカル表示が、次のウィンドウに表示されます。

### Amazon SQS

Amazon SQS 通知を有効にすると、EMRFS の初期化時に EMRFS-Inconsistency-<jobFlowId> という名前の Amazon SQS キューが作成されます。Amazon S3 の結果整合性のために FileSystem API 呼び出しが失敗すると、Amazon SQS メッセージがキューにプッシュされます。メッセージには JobFlowId、API、整合性のないパスのリスト、スタックトレースなどの情報が含まれます。メッセージは、Amazon SQS コンソールを使用するか、EMRFS の read-sqs コマンドを使用して読み取ることができます。

## Amazon S3 の結果整合性の問題に対する Amazon SQS メッセージを管理するには

Amazon S3 の結果整合性の問題に関する Amazon SQS メッセージは、EMRFS CLI を使用して読み取ることができます。Amazon SQS EMRFS キューからメッセージを読み取るには、`read-sqs` コマンドを入力し、マスターノードのローカルファイルシステム上の結果の出力ファイルの出力場所を指定します。

`delete-sqs` コマンドを使用して EMRFS の Amazon SQS キューを削除することもできます。

1. Amazon SQS キューからメッセージを読み取るには、次のコマンドを入力します。*queuename* を設定した Amazon SQS キューの名前に、*/path/filename* を出力ファイルのパスに置き換えてください。

```
emrfs read-sqs --queue-name queuename --output-file /path/filename
```

たとえば、デフォルトのキューから Amazon SQS メッセージを読み取って出力するには、次のように入力します。

```
emrfs read-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU --output-file /path/filename
```

### Note

また、`-q` および `-o` の代わりに、それぞれ `--queue-name` および `--output-file` ショートカットを使用することもできます。

2. Amazon SQS キューを削除するには、次のコマンドを入力します。

```
emrfs delete-sqs --queue-name queuename
```

たとえば、デフォルトキューを削除するには、次のように入力します。

```
emrfs delete-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU
```

### Note

`--queue-name` の代わりに `-q` ショートカットを使用することもできます。

## 整合性のあるビューを設定する

emrfs-site プロパティの設定プロパティを使用して、整合性のあるビューに追加の設定を構成できます。たとえば、emrfs-site 設定分類 (Amazon EMR リリースバージョン 4.x 以降のみ)、またはマスターノードに emrfs-site.xml ファイルを設定するブートストラップアクションを使用して、CLI の `--emrfs` オプションに次の引数を指定することで、デフォルトの DynamoDB スループットとして異なる値を選択できます。

Example クラスター起動時のデフォルトのメタデータの読み取りおよび書き込み値の変更

```
aws emr create-cluster --release-label emr-7.1.0 --instance-type m5.xlarge \  
--emrfs Consistent=true,Args=[fs.s3.consistent.metadata.read.capacity=600,\  
fs.s3.consistent.metadata.write.capacity=300] --ec2-attributes KeyName=myKey
```

別の方法として、次の設定ファイルを使用し、ローカルまたは Amazon S3 に保存します。

```
[  
  {  
    "Classification": "emrfs-site",  
    "Properties": {  
      "fs.s3.consistent.metadata.read.capacity": "600",  
      "fs.s3.consistent.metadata.write.capacity": "300"  
    }  
  }  
]
```

次の構文で作成した設定を使用します。

```
aws emr create-cluster --release-label emr-7.1.0 --applications Name=Hive \  
--instance-type m5.xlarge --instance-count 2 --configurations file:///./myConfig.json
```

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

以下のオプションは、設定または AWS CLI `--emrfs` 引数を使用して設定できます。それらの引数については、「[AWS CLI コマンドリファレンス](#)」を参照してください。

整合性のあるビューの `emrfs-site.xml` プロパティ

プロパティ	デフォルト値	説明
<code>fs.s3.consistent</code>	<b>false</b>	<b>true</b> に設定すると、このプロパティは DynamoDB を使用して整合性を確保するように EMRFS を設定します。
<code>fs.s3.consistent.retryPolicyType</code>	<b>exponential</b>	このプロパティは、整合性に関する問題について再試行するときに使用するポリシーを識別します。オプションは、 <code>exponential</code> 、 <code>fixed</code> 、または <code>none</code> です。
<code>fs.s3.consistent.retryPeriodSeconds</code>	<b>1</b>	このプロパティは、整合性の再試行間の待機時間を設定します。
<code>fs.s3.consistent.retryCount</code>	<b>10</b>	このプロパティは、不整合が検出されたときの最大の再試行回数を設定します。
<code>fs.s3.consistent.throwExceptionOnInconsistency</code>	<b>true</b>	このプロパティは、整合性の例外をスローまたは記録するかどうかを決定します。 <b>true</b> に設定すると、 <code>ConsistencyException</code> がスローされます。
<code>fs.s3.consistent.metadata.autoCreate</code>	<b>true</b>	<b>true</b> に設定すると、このプロパティはメタデータテーブルの自動作成を有効にします。
<code>fs.s3.consistent.metadata.tag.verification.enabled</code>	<b>true</b>	Amazon EMR 5.29.0 では、このプロパティはデフォルトで有

プロパティ	デフォルト値	説明
		効になっています。有効にすると、EMRFS は S3 ETags を使用して、読み取られているオブジェクトが利用可能な最新バージョンであることを確認します。この機能は、同じ名前を保持したまま S3 上のファイルが上書きされる read-after-update ユースケースに役立ちます。この ETag 検証機能は、現在 S3 Select では使用できません。
<code>fs.s3.consistent.metadata.tableName</code>	<b>EmrFSMetadata</b>	このプロパティは、DynamoDB 内のメタデータテーブルの名前を指定します。
<code>fs.s3.consistent.metadata.read.capacity</code>	<b>500</b>	このプロパティは、メタデータテーブルの作成時に DynamoDB の読み込みキャパシティーをプロビジョニングすることを指定します。
<code>fs.s3.consistent.metadata.write.capacity</code>	<b>100</b>	このプロパティは、メタデータテーブルの作成時に DynamoDB の書き込みキャパシティーをプロビジョニングすることを指定します。
<code>fs.s3.consistent.fastList</code>	<b>true</b>	<b>true</b> に設定すると、このプロパティは複数のスレッドを使用してディレクトリを一覧表示します (必要な場合)。このプロパティを使用するには、整合性を有効にする必要があります。

プロパティ	デフォルト値	説明
<code>fs.s3.consistent.fastList.p refetchMetadata</code>	<b>false</b>	<b>true</b> に設定すると、このプロパティは 20,000 以上の項目を含むディレクトリに対してメタデータのプリフェッチを有効にします。
<code>fs.s3.consistent.notification.CloudWatch</code>	<b>false</b>	に設定すると <b>true</b> 、Amazon S3 の結果整合性の問題により失敗した FileSystem API コールに対して CloudWatch メトリクスが有効になります。Amazon S3
<code>fs.s3.consistent.notification.SQS</code>	<b>false</b>	<b>true</b> に設定すると、結果整合性の通知が Amazon SQS キューにプッシュされます。
<code>fs.s3.consistent.notification.SQS.queueName</code>	<b>EMRFS-Inconsistent- &lt;jobFlowId&gt;</b>	このプロパティを変更すると、Amazon S3 の結果整合性の問題に関するメッセージに対して、独自の SQS キュー名を指定できます。
<code>fs.s3.consistent.notification.SQS.customMsg</code>	<b>none</b>	このプロパティでは、Amazon S3 の結果整合性の問題に関する SQS メッセージに含まれるカスタム情報を指定できます。このプロパティに対して値が指定されていない場合、メッセージの対応するフィールドは空になります。
<code>fs.s3.consistent.dynamodb.endpoint</code>	<b>none</b>	このプロパティを使用して、整合性のあるビューのメタデータにカスタム DynamoDB エンドポイントを指定できます。

プロパティ	デフォルト値	説明
fs.s3.useRequesterPaysHeader	<b>false</b>	<b>true</b> に設定すると、このプロパティは、リクエスト支払いオプションが有効になっているバケットへの Amazon S3 リクエストを許可します。

## EMRFS CLI コマンドリファレンス

EMRFS CLI は、Amazon EMR リリースバージョン 3.2.1 以降を使用して作成されたすべてのクラスターマスターノードにデフォルトでインストールされます。EMRFS CLI を使用して、整合性のあるビューのメタデータを管理できます。

### Note

emrfs コマンドは、VT100 ターミナルエミュレーションでのみサポートされます。ただし、他のターミナルエミュレータモードでも動作する場合があります。

## emrfs 最上位コマンド

emrfs 最上位コマンドは、次の構造をサポートしています。

```
emrfs [describe-metadata | set-metadata-capacity | delete-metadata | create-metadata |
\
list-metadata-stores | diff | delete | sync | import ] [options] [arguments]
```

次の表に示すように、[引数] の有無にかかわらず、[options] を指定します。サブコマンド (describe-metadata、set-metadata-capacity など) に固有の [options] については、以下の各サブコマンドを参照してください。

### emrfs の [Options]

オプション	説明	必須
-------	----	----



オプション	説明	必須
-a <i>AWS_ACCESS_KEY_ID</i>   --access-key <i>AWS_ACCESS_KEY_ID</i>	Amazon S3 にオブジェクトを書き込み、DynamoDB でメタデータストアを作成またはアクセスするために使用する AWS アクセスキー。デフォルトでは、 <i>AWS_ACCESS_KEY_ID</i> は、クラスターの作成に使用されたアクセスキーに設定されます。	いいえ
-s <i>AWS_SECRET_ACCESS_KEY</i>   --secret-key <i>AWS_SECRET_ACCESS_KEY</i>	Amazon S3 へのオブジェクトの書き込みと、DynamoDB のメタデータストアの作成またはアクセスに使用するアクセスキーに関連付けられた AWS シークレットキー。DynamoDB デフォルトでは、 <i>AWS_SECRET_ACCESS_KEY</i> は、クラスターの作成に使用されたアクセスキーに関連付けられたシークレットキーに設定されます。	いいえ
-v   --verbose	詳細な出力を作成します。	いいえ
-h   --help	emrfs コマンドのヘルプメッセージと使用方法の説明を表示します。	いいえ

## emrfs describe-metadata サブコマンド

### emrfs describe-metadata の [Options]

オプション	説明	必須
-m <i>METADATA_NAME</i>   --metadata-name <i>METADATA_NAME</i>	<i>METADATA_NAME</i> は DynamoDB メタデータテーブルの名前です。 <i>METADATA_NAME</i> 引数が指定されていない場合、デフォルト値は EmrFSMetadata です。	いいえ

## Example emrfs describe-metadata の例

次の例では、デフォルトのメタデータテーブルを記述します。

```
$ emrfs describe-metadata
EmrFSMetadata
  read-capacity: 400
  write-capacity: 100
  status: ACTIVE
  approximate-item-count (6 hour delay): 12
```

## emrfs set-metadata-capacity サブコマンド

emrfs の [オプション] set-metadata-capacity

オプション	説明	必須
-m <i>METADATA_NAME</i>   --metadata-name <i>METADATA_NAME</i>	<i>METADATA_NAME</i> は DynamoDB メタデータテーブルの名前です。 <i>METADATA_NAME</i> 引数が指定されていない場合、デフォルト値は EmrFSMetadata です。	いいえ
-r <i>READ_CAPACITY</i>   --read-capacity <i>READ_CAPACITY</i>	メタデータテーブルのリクエストされた読み込みスループット容量。 <i>READ_CAPACITY</i> 引数が指定されていない場合、デフォルト値は 400 です。	いいえ
-w <i>WRITE_CAPACITY</i>   --write-capacity <i>WRITE_CAPACITY</i>	メタデータテーブルのリクエストされた書き込みスループット容量。 <i>WRITE_CAPACITY</i> 引数が指定されていない場合、デフォルト値は 100 です。	いいえ

## Example emrfs set-metadata-capacity の例

次の例では、600 という名前のメタデータテーブルの読み込みスループット容量を 150、書き込みスループット容量を EmrMetadataAlt に設定します。

```
$ emrfs set-metadata-capacity --metadata-name EmrMetadataAlt --read-capacity 600 --write-capacity 150
```

```
read-capacity: 400
write-capacity: 100
status: UPDATING
approximate-item-count (6 hour delay): 0
```

## emrfs delete-metadata サブコマンド

### emrfs delete-metadata の [Options]

オプション	説明	必須
-m <i>METADATA_NAME</i>   --metadata-name <i>METADATA_NAME</i>	<i>METADATA_NAME</i> は DynamoDB メタデータテーブルの名前です。 <i>METADATA_NAME</i> 引数が指定されていない場合、デフォルト値は EmrFSMeta data です。	いいえ

### Example emrfs delete-metadata の例

次の例では、デフォルトのメタデータテーブルを削除します。

```
$ emrfs delete-metadata
```

## emrfs create-metadata サブコマンド

### emrfs create-metadata の [Options]

オプション	説明	必須
-m <i>METADATA_NAME</i>   --metadata-name <i>METADATA_NAME</i>	<i>METADATA_NAME</i> は DynamoDB メタデータテーブルの名前です。 <i>METADATA_NAME</i> 引数が指定されていない場合、デフォルト値は EmrFSMeta data です。	いいえ
-r <i>READ_CAPACITY</i>   --read-capacity <i>READ_CAPACITY</i>	メタデータテーブルのリクエストされた読み込みスループット容量。 <i>READ_CAPACITY</i> 引数が指定されていない場合、デフォルト値は 400 です。	いいえ

オプション	説明	必須
-w <i>WRITE_CAPACITY</i>   --write-capacity <i>WRITE_CAPACITY</i>	メタデータテーブルのリクエストされた書き込みスループット容量。 <i>WRITE_CAPACITY</i> 引数が指定されていない場合、デフォルト値は 100 です。	いいえ

### Example emrfs create-metadata の例

次の例では、EmrFSMetadataAlt という名前のメタデータテーブルを作成します。

```
$ emrfs create-metadata -m EmrFSMetadataAlt
Creating metadata: EmrFSMetadataAlt
EmrFSMetadataAlt
  read-capacity: 400
  write-capacity: 100
  status: ACTIVE
  approximate-item-count (6 hour delay): 0
```

### emrfs list-metadata-stores サブコマンド

emrfs list-metadata-stores サブコマンドには、[options] がありません。

### Example L ist-metadata-stores の例

次の例では、メタデータテーブルを表示します。

```
$ emrfs list-metadata-stores
EmrFSMetadata
```

### emrfs diff サブコマンド

emrfs diff の [Options]

オプション	説明	必須
-m <i>METADATA_NAME</i>   --metadata-name <i>METADATA_NAME</i>	<i>METADATA_NAME</i> は DynamoDB メタデータテーブルの名前です。 <i>METADATA_NAME</i> 引数が指定されていない場合、デフォルト値は EmrFSMetadata です。	いいえ

オプション	説明	必須
<code>s3://s3Path</code>	メタデータテーブルと比較する Amazon S3 バケットへのパス。バケットは再帰的に同期されます。	はい

## Example emrfs diff の例

次の例では、デフォルトのメタデータテーブルを Amazon S3 バケットと比較します。

```
$ emrfs diff s3://elasticmapreduce/samples/cloudfront
BOTH | MANIFEST ONLY | S3 ONLY
DIR elasticmapreduce/samples/cloudfront
DIR elasticmapreduce/samples/cloudfront/code/
DIR elasticmapreduce/samples/cloudfront/input/
DIR elasticmapreduce/samples/cloudfront/logprocessor.jar
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-14.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-15.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-16.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-17.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-18.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-19.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-20.WxYz1234
DIR elasticmapreduce/samples/cloudfront/code/cloudfront-loganalyzer.tgz
```

## emrfs delete サブコマンド

### emrfs delete の [Options]

オプション	説明	必須
<code>-m METADATA_NAME</code>   <code>--metadata-name METADATA_NAME</code>	<code>METADATA_NAME</code> は DynamoDB メタデータテーブルの名前です。 <code>METADATA_NAME</code> 引数が指定されていない場合、デフォルト値は <code>EmrFSMetadata</code> です。	いいえ
<code>s3://s3Path</code>	整合性のあるビューのために追跡する Amazon S3 バケットへのパス。バケットは再帰的に同期されます。	はい

オプション	説明	必須
<code>-t <i>TIME</i>   --time <i>TIME</i></code>	有効期限 ( 時間単位引数を使用して解釈されます )。指定されたバケットの <i>TIME</i> 引数より古いすべてのメタデータエントリが削除されます。	
<code>-u <i>UNIT</i>   --time-unit <i>UNIT</i></code>	時間の引数 ( ナノ秒、マイクロ秒、ミリ秒、秒、分、時間、日数 ) を解釈するために使用される基準。引数が指定されていない場合、デフォルト値は days です。	
<code>--read-consumption <i>READ_CONSUMPTION</i></code>	delete オペレーションに使用される、リクエストされた利用可能な読み取りスループットの量。 <i>READ_CONSUMPTION</i> 引数が指定されていない場合、デフォルト値は 400 です。	いいえ
<code>--write-consumption <i>WRITE_CONSUMPTION</i></code>	delete オペレーションに使用される、リクエストされた利用可能な書き込みスループットの量。 <i>WRITE_CONSUMPTION</i> 引数が指定されていない場合、デフォルト値は 100 です。	いいえ

### Example emrfs delete の例

次の例では、整合性のあるビューの追跡メタデータから Amazon S3 バケットのすべてのオブジェクトを削除します。

```
$ emrfs delete s3://elasticmapreduce/samples/cloudfront
entries deleted: 11
```

### emrfs import サブコマンド

#### emrfs import の [Options]

オプション	説明	必須
-------	----	----

オプション	説明	必須
<code>-m <i>METADATA_NAME</i></code>   <code>--metadata-name <i>METADATA_NAME</i></code>	<i>METADATA_NAME</i> は DynamoDB メタデータテーブルの名前です。 <i>METADATA_NAME</i> 引数が指定されていない場合、デフォルト値は <code>EmrFSMetadata</code> です。	いいえ
<code>s3://s3Path</code>	整合性のあるビューのために追跡する Amazon S3 バケットへのパス。バケットは再帰的に同期されます。	はい
<code>--read-consumption <i>READ_CONSUMPTION</i></code>	delete オペレーションに使用される、リクエストされた利用可能な読み取りスループットの量。 <i>READ_CONSUMPTION</i> 引数が指定されていない場合、デフォルト値は 400 です。	いいえ
<code>--write-consumption <i>WRITE_CONSUMPTION</i></code>	delete オペレーションに使用される、リクエストされた利用可能な書き込みスループットの量。 <i>WRITE_CONSUMPTION</i> 引数が指定されていない場合、デフォルト値は 100 です。	いいえ

### Example emrfs import の例

次の例では、Amazon S3 バケット内のすべてのオブジェクトを、整合性のあるビューの追跡メタデータとともにインポートします。不明なキーはすべて無視されます。

```
$ emrfs import s3://elasticmapreduce/samples/cloudfront
```

### emrfs sync サブコマンド

emrfs sync の [Options]

オプション	説明	必須
-------	----	----

オプション	説明	必須
<code>-m <i>METADATA_NAME</i></code>   <code>--metadata-name</code> <i>METADATA_NAME</i>	<i>METADATA_NAME</i> は DynamoDB メタデータテーブルの名前です。 <i>METADATA_NAME</i> 引数が指定されていない場合、デフォルト値は <code>EmrFSMetadata</code> です。	いいえ
<i>s3://s3Path</i>	整合性のあるビューのために追跡する Amazon S3 バケットへのパス。バケットは再帰的に同期されます。	はい
<code>--read-consumption</code> <i>READ_CONSUMPTION</i>	<code>delete</code> オペレーションに使用される、リクエストされた利用可能な読み取りスループットの量。 <i>READ_CONSUMPTION</i> 引数が指定されていない場合、デフォルト値は 400 です。	いいえ
<code>--write-consumption</code> <i>WRITE_CONSUMPTION</i>	<code>delete</code> オペレーションに使用される、リクエストされた利用可能な書き込みスループットの量。 <i>WRITE_CONSUMPTION</i> 引数が指定されていない場合、デフォルト値は 100 です。	いいえ

### Example emrfs sync command の例

次の例では、Amazon S3 バケット内のすべてのオブジェクトを、整合性のあるビューの追跡メタデータとともにインポートします。不明なキーがすべて削除されます。

```
$ emrfs sync s3://elasticmapreduce/samples/cloudfront
Synching samples/cloudfront                0 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/code/          1 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/               2 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/input/         9 added | 0 updated |
 0 removed | 0 unchanged
Done synching s3://elasticmapreduce/samples/cloudfront
1 removed | 0 unchanged
```



```
creating 3 folder key(s)
folders written: 3
```

## emrfs read-sqs サブコマンド

### emrfs read-sqs の [Options]

オプション	説明	必須
<code>-q <i>QUEUE_NAME</i>   --queue-name <i>QUEUE_NAME</i></code>	<i>QUEUE_NAME</i> は、emrfs-site.xml で設定した Amazon SQS キューの名前です。デフォルト値は、 <b>EMRFS-Inconsistency-&lt;jobFlowId&gt;</b> です。	はい
<code>-o <i>OUTPUT_FILE</i>   --output-file <i>OUTPUT_FILE</i></code>	<i>OUTPUT_FILE</i> は、マスターノードのローカルファイルシステムにある出力ファイルへのパスです。キューから読み取ったメッセージはこのファイルに書き込まれます。	はい

## emrfs delete-sqs サブコマンド

### emrfs delete-sqs の [Options]

オプション	説明	必須
<code>-q <i>QUEUE_NAME</i>   --queue-name <i>QUEUE_NAME</i></code>	<i>QUEUE_NAME</i> は、emrfs-site.xml で設定した Amazon SQS キューの名前です。デフォルト値は、 <b>EMRFS-Inconsistency-&lt;jobFlowId&gt;</b> です。	はい

## ステップとして EMRFS CLI コマンドを送信する

次の例は、AWS CLI または API を活用してマスターノードで emrfs ユーティリティを使用し `command-runner.jar`、emrfs コマンドをステップとして実行する方法を示しています。この例では AWS SDK for Python (Boto3) を使用して、Amazon S3 バケット内のオブジェクトをデフォルトの EMRFS メタデータテーブルに追加するステップをクラスターに追加します。

```
import boto3
from botocore.exceptions import ClientError

def add_emrfs_step(command, bucket_url, cluster_id, emr_client):
    """
    Add an EMRFS command as a job flow step to an existing cluster.

    :param command: The EMRFS command to run.
    :param bucket_url: The URL of a bucket that contains tracking metadata.
    :param cluster_id: The ID of the cluster to update.
    :param emr_client: The Boto3 Amazon EMR client object.
    :return: The ID of the added job flow step. Status can be tracked by calling
             the emr_client.describe_step() function.
    """
    job_flow_step = {
        "Name": "Example EMRFS Command Step",
        "ActionOnFailure": "CONTINUE",
        "HadoopJarStep": {
            "Jar": "command-runner.jar",
            "Args": ["/usr/bin/emrfs", command, bucket_url],
        },
    },
}

    try:
        response = emr_client.add_job_flow_steps(
            JobFlowId=cluster_id, Steps=[job_flow_step]
        )
        step_id = response["StepIds"][0]
        print(f"Added step {step_id} to cluster {cluster_id}.")
    except ClientError:
        print(f"Couldn't add a step to cluster {cluster_id}.")
        raise
    else:
        return step_id

def usage_demo():
    emr_client = boto3.client("emr")
    # Assumes the first waiting cluster has EMRFS enabled and has created metadata
    # with the default name of 'EmrFSMetadata'.
    cluster = emr_client.list_clusters(ClusterStates=["WAITING"])["Clusters"][0]
    add_emrfs_step(
```

```
        "sync", "s3://elasticmapreduce/samples/cloudfront", cluster["Id"], emr_client
    )

if __name__ == "__main__":
    usage_demo()
```

返された `step_id` 値を使用して、ログで操作の結果を確認できます。

## Amazon S3 内の EMRFS データへのアクセスを許可する

デフォルトでは、EC2 の EMR ロールが、Amazon S3 内の EMRFS データにアクセスする権限を決定します。EMRFS 経由でリクエストを行うユーザーまたはグループに関係なく、このロールにアタッチされている IAM ポリシーが適用されます。デフォルトは `EMR_EC2_DefaultRole` です。詳細については、「[クラスター EC2 インスタンスのサービスロール \(EC2 インスタンスプロファイル\)](#)」を参照してください。

Amazon EMR リリースバージョン 5.10.0 からは、セキュリティ設定を使用して、EMRFS の IAM ロールを指定できます。これにより、複数のユーザーを持つクラスターについて、Amazon S3 への EMRFS リクエストの権限をカスタマイズできます。Amazon S3 内のプレフィックスに基づいて、異なるユーザーとグループ、および異なる Amazon S3 バケットの場所に対して、異なる IAM ロールを指定できます。EMRFS が Amazon S3 に対して、指定したユーザー、グループ、または場所と一致するリクエストを行うと、クラスターは EC2 の EMR ロールの代わりに、ユーザーが指定した対応するロールを使用します。詳細については、「[Amazon S3 への EMRFS リクエストの IAM ロールを設定する](#)」を参照してください。

また、Amazon EMR ソリューションの要求が EMRFS の IAM ロールで提供される権限を超える場合、カスタムの認証情報プロバイダークラスを定義して Amazon S3 内の EMRFS データへのアクセス権限をカスタマイズできます。

### Amazon S3 内の EMRFS データに対してカスタムの認証情報プロバイダーを作成する

カスタム認証情報プロバイダーを作成するには、[AWSCredentialsProvider](#) と Hadoop [設定可能](#) クラスを実装します。

このアプローチの詳細な説明については、ビッグデータブログの「[EMRFS を使用して別の AWS アカウントのデータを安全に分析する AWS](#)」を参照してください。ブログ記事には、IAM ロールの作

成からクラスターの起動までend-to-end、プロセスを説明するチュートリアルが含まれています。また、カスタム認証情報プロバイダークラスを実装する Java コードの例も提供しています。

基本的なステップは次のとおりです。

カスタム認証情報プロバイダーを指定するには

1. JAR ファイルとしてコンパイルされる、カスタム認証情報プロバイダークラスを作成します。
2. ブートストラップアクションとしてスクリプトを実行し、カスタム認証情報プロバイダーの JAR ファイルを、クラスターのマスターノードの `/usr/share/aws/emr/emrfs/auxlib` の場所にコピーします。ブートストラップアクションの詳細については、「[\(オプション\) 追加のソフトウェアをインストールするためのブートストラップアクションの作成](#)」を参照してください。
3. `emrfs-site` 分類をカスタマイズして、JAR ファイルで実装するクラスを指定します。アプリケーションをカスタマイズするための設定オブジェクトの指定に関する詳細は、「Amazon EMR リリース ガイド」の「[アプリケーションの設定](#)」を参照してください。

次の例では、共通設定パラメータを使って Hive クラスターを起動する `create-cluster` コマンドを示しています。さらに次のものが含まれています。

- Amazon S3 の `mybucket` に保存されるスクリプト `copy_jar_file.sh` を実行するブートストラップアクション。
- JAR ファイルで `emrfs-site` として定義されたカスタム認証情報プロバイダーを指定する `MyCustomCredentialsProvider` 分類

#### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --applications Name=Hive \  
--bootstrap-actions '[{"Path":"s3://mybucket/copy_jar_file.sh","Name":"Custom action"}]' \  
--ec2-attributes '{"KeyName":"MyKeyPair","InstanceProfile":"EMR_EC2_DefaultRole",\  
"SubnetId":"subnet-xxxxxxxx","EmrManagedSlaveSecurityGroup":"sg-xxxxxxxx",\  

```

```
"EmrManagedMasterSecurityGroup":"sg-xxxxxxx"}' \  
--service-role EMR_DefaultRole_V2 --enable-debugging --release-label emr-7.1.0 \  
--log-uri 's3n://my-emr-log-bucket/' --name 'test-awscredentialsprovider-emrfs' \  
--instance-type=m5.xlarge --instance-count 3 \  
--configurations '[{"Classification":"emrfs-site",\  
"Properties":\  
{  
"fs.s3.customAWSCredentialsProvider":"MyAWSCredentialsProviderWithUri"},\  
"Configurations":[]}]'
```

## デフォルト AWS Security Token Service エンドポイントの管理

EMRFS は、AWS Security Token Service (STS) を使用して一時的なセキュリティ認証情報を取得し、AWS リソースにアクセスします。以前の Amazon EMR リリースバージョンでは、すべての AWS STS リクエストがの単一のグローバルエンドポイントに送信されます `https://sts.amazonaws.com`。Amazon EMR リリースバージョン 5.31.0 および 6.1.0 以降では、代わりにリージョン AWS STS エンドポイントにリクエストを行います。これにより、待ち時間が短縮され、セッショントークンの有効性が向上します。エンドポイントの詳細については、AWS STS 「[ユーザーガイドAWS STS](#)」の [AWS 「リージョンでの管理](#)」AWS Identity and Access Management」を参照してください。

Amazon EMR リリースバージョン 5.31.0 と 6.1.0 以降を使用すると、デフォルトの AWS STS エンドポイントを上書きできます。そのためには、`emrfs-site` 設定内の `fs.s3.sts.endpoint` プロパティを変更する必要があります。

次の AWS CLI 例では、EMRFS で使用されるデフォルトの AWS STS エンドポイントをグローバルエンドポイントに設定します。

```
aws emr create-cluster --release-label <emr-5.33.0> --instance-type m5.xlarge \  
--emrfs Args=[fs.s3.sts.endpoint=https://sts.amazonaws.com]
```

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

別の方法として、次の例を使用して JSON 設定ファイルを作成し、`emr create-cluster` の `--configurations` 引数を使用してそのファイルを指定します。`--configurations` の使用に関する詳細は、「[AWS CLI コマンドリファレンス](#)」を参照してください。

```
[
  {
    "classification": "emrfs-site",
    "properties": {
      "fs.s3.sts.endpoint": "https://sts.amazonaws.com"
    }
  }
]
```

## EMRFS プロパティを使用して Amazon S3 の暗号化を指定する

### Important

Amazon EMR リリースバージョン 4.8.0 から、セキュリティ設定を使用して、より簡単により豊富なオプションで暗号化設定を適用できるようになりました。セキュリティ設定を使用することをお勧めします。詳細については、「[データ暗号化の設定](#)」を参照してください。このセクションで説明するコンソールの手順は、リリース 4.8.0 より前のリリースバージョンで使用できます。を使用してクラスター設定と後続バージョンのセキュリティ設定の両方で Amazon S3 暗号化 AWS CLI を設定する場合、セキュリティ設定はクラスター設定を上書きします。

クラスターを作成するときは、コンソールを使用するか、または EMR SDK で `emrfs-site` 分類プロパティを使用して、Amazon S3 の EMRFS データのサーバー側の暗号化 (SSE) AWS CLI またはクライアント側の暗号化 (CSE) を指定できます。Amazon S3 の SSE と CSE は相互排他的であるため、いずれかを選択できますが、両方は選択できません。

AWS CLI 手順については、以下の暗号化タイプに適したセクションを参照してください。

を使用して EMRFS 暗号化オプションを指定するには AWS Management Console

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。

2. [Create cluster (クラスターの作成)], [Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. 4.7.2 以前の [Release] を選択します。
4. アプリケーションに合わせて [Software and Steps] の他のオプションを選択し、[Next] を選択します。
5. アプリケーションに合わせて、[Hardware] ペインおよび [General Cluster Settings] ペインで設定を選択します。
6. [Security] ペインの [Authentication and encryption] で、使用する [S3 Encryption (with EMRFS)] オプションを選択します。

#### Note

Amazon EMR リリースバージョン 4.4 以前を使用しているときは、[KMS Key Management を使用した S3 サーバー側の暗号化] (SSE-KMS) を使用できません。

- AWS Key Management を使用するオプションを選択する場合は、[AWS KMS キー ID] を選択します。詳細については、「[EMRFS 暗号化 AWS KMS keys に を使用する](#)」を参照してください。
  - [S3 client-side encryption with custom materials provider] を選択した場合、[Class name] と [JAR location] に入力してください。詳細については、「[Amazon S3 クライアント側の暗号化](#)」を参照してください。
7. 必要に応じて、アプリケーションの他のオプションを選択し、[Create Cluster] を選択します。

## EMRFS 暗号化 AWS KMS keys に を使用する

AWS KMS 暗号化キーは、Amazon EMR クラスターインスタンスおよび EMRFS で使用される Amazon S3 バケットと同じリージョンで作成する必要があります。指定するキーが、クラスターの設定に使用するアカウントとは異なるアカウントにある場合は、その ARN を使用してキーを指定する必要があります。

Amazon EC2 インスタンスプロファイルのロールには、指定した KMS キーを使用するためのアクセス許可が必要です。Amazon EMR 内のインスタンスプロファイルのデフォルトのロールは EMR\_EC2\_DefaultRole です。インスタンスプロファイルに別のロールを使用する場合、または Amazon S3 への EMRFS リクエストに IAM ロールを使用する場合は、必要に応じて各ロールがキーユーザーとして追加されていることを確認してください。これにより、KMS キーを使用するアクセ



ス許可がロールに付与されます。詳細については、「AWS Key Management Service デベロッパーガイド」の「[キーポリシーの使用](#)」と、「[Amazon S3 への EMRFS リクエストの IAM ロールの設定](#)」を参照してください。

を使用して AWS Management Console、指定した KMS キーのキーユーザーのリストにインスタンスプロファイルまたは EC2 インスタンスプロファイルを追加するか、AWS CLI または AWS SDK を使用して適切なキーポリシーをアタッチできます。

Amazon EMR は、[対称 KMS キー](#)のみをサポートします。[非対称 KMS キー](#)を使用して、Amazon EMR クラスター内の保管中のデータを暗号化することはできません。KMS キーが対称か非対称かを判別するには、「[対称および非対称 KMS キーを識別する](#)」を参照してください。

以下の手順では、AWS Management Consoleを使用して、デフォルトの Amazon EMR インスタンスプロファイル EMR\_EC2\_DefaultRole をキーユーザーとして追加する方法について説明します。既に KMS キーが作成されていることを前提としています。新しい KMS キーを作成するには、「AWS Key Management Service デベロッパーガイド」の「[キーの作成](#)」を参照してください。

暗号化キーユーザーのリストに Amazon EMR の EC2 インスタンスプロファイルを追加するには

1. にサインイン AWS Management Console し、<https://console.aws.amazon.com/kms> で AWS Key Management Service (AWS KMS) コンソールを開きます。
2. を変更するには AWS リージョン、ページの右上隅にあるリージョンセレクタを使用します。
3. 変更する KMS キーのエイリアスを選択します。
4. [Key Users] のキーの詳細ページで、[Add] を選択します。
5. [Add key users] ダイアログボックスで、適切なロールを選択します。デフォルトロールの名前は EMR\_EC2\_DefaultRole です。
6. [追加] を選択します。

## Amazon S3 のサーバー側の暗号化

Amazon S3 のサーバー側の暗号化をセットアップすると、Amazon S3 はデータをディスクに書き込むときにオブジェクトレベルで暗号化し、アクセスするときに復号します。SSE に関する詳細は、「Amazon Simple Storage Service ユーザーガイド」の「[サーバー側の暗号化を使用したデータの保護](#)」を参照してください。

Amazon EMR で SSE を指定するときに、次の 2 つの異なるキー管理システムから選択できます。

- SSE-S3 — Amazon S3 がキーを管理します。



- SSE-KMS – を使用して AWS KMS key、Amazon EMR に適したポリシーをセットアップします。Amazon EMR のキー要件の詳細については、[「暗号化に使用する AWS KMS keys」](#)を参照してください。

SSE とお客様が用意したキーとの組み合わせ (SSE-C) は、Amazon EMR では使用できません。

を使用して SSE-S3 を有効にしたクラスターを作成するには AWS CLI

- 次のコマンドを入力します。

```
aws emr create-cluster --release-label emr-4.7.2 or earlier \  
--instance-count 3 --instance-type m5.xlarge --emrfs Encryption=ServerSide
```

emrfs-site プロパティで fs.s3.Encryption プロパティを true に設定することで、SSE-S3 を有効にすることもできます。enableServerSide以下の SSE-KMS の例を参照して、キー ID のプロパティを省略します。

を使用して SSE-KMS を有効にしたクラスターを作成するには AWS CLI

#### Note

SSE-KMS は、Amazon EMR リリースバージョン 4.5.0 以降でのみ使用できます。

- 次の AWS CLI コマンドを入力して、SSE-KMS でクラスターを作成します。*keyID* は AWS KMS key です。例えば、*a4567b8-9900-12ab-1234-123a45678901* です。

```
aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3 \  
--instance-type m5.xlarge --use-default-roles \  
--emrfs Encryption=ServerSide,Args=[fs.s3.serverSideEncryption.kms.keyId=keyId]
```

--または--

emrfs-site 分類を使用して次の AWS CLI コマンドを入力し、次のmyConfig.json例のようにコンテンツを含む設定 JSON ファイルを指定します。

```
aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3
--instance-type m5.xlarge --applications Name=Hadoop --configurations file://
myConfig.json --use-default-roles
```

myConfig.json のサンプルコンテンツ:

```
[
  {
    "Classification": "emrfs-site",
    "Properties": {
      "fs.s3.enableServerSideEncryption": "true",
      "fs.s3.serverSideEncryption.kms.keyId": "a4567b8-9900-12ab-1234-123a45678901"
    }
  }
]
```

## SSE-S3 と SSE-KMS の設定プロパティ

これらのプロパティは `emrfs-site` 設定分類を使用して設定できます。SSE-KMS は、Amazon EMR リリースバージョン 4.5.0 以降でのみ使用できます。

プロパティ	デフォルト値	説明
<code>fs.s3.enableServerSideEncryption</code>	<b>false</b>	<b>true</b> に設定すると、Amazon S3 に保存されたオブジェクトはサーバー側の暗号化を使用して暗号化されます。キーが指定されていない場合、SSE-S3 が使用されます。
<code>fs.s3.serverSideEncryption.kms.keyId</code>	<b>n/a</b>	AWS KMS キー ID または ARN を指定します。キーを指定した場合、SSE-KMS が使用されます。

## Amazon S3 クライアント側の暗号化

Amazon S3 のクライアント側の暗号化を使用すると、Amazon S3 の暗号化と復号はクラスターの EMRFS クライアントで行われます。オブジェクトは Amazon S3 にアップロードされる前に暗号化され、ダウンロード後に復号化されます。指定するプロバイダーが、クライアントが使用する暗号化キーを提供します。クライアントは、AWS KMS によって提供されるキー (CSE-KMS) か、クライアント側のルートキーを提供するカスタム Java クラス (CSE-C) を使用できます。CSE-KMS と CSE-C では、指定するプロバイダーと、復号化または暗号化されるオブジェクトのメタデータに応じて、暗号化の仕様が少し異なります。これらの差異に関する詳細は、「Amazon Simple Storage Service ユーザーガイド」の「[クライアント側の暗号化を使用したデータの保護](#)」を参照してください。

### Note

Amazon S3 CSE では、Amazon S3 と交換される EMRFS データのみが暗号化されます。クラスターインスタンスボリュームのすべてのデータが暗号化されるわけではありません。さらに、Hue は EMRFS を使用しないため、Hue S3 ファイルブラウザが Amazon S3 に書き込むオブジェクトは暗号化されません。

を使用して Amazon S3 の EMRFS データの CSE-KMS を指定するには AWS CLI

- 次のコマンドを入力して、*MyKMSKeyID* を、使用する KMS キーのキー ID または ARN に置き換えます。

```
aws emr create-cluster --release-label emr-4.7.2 or earlier
--emrfs Encryption=ClientSide,ProviderType=KMS,KMSKeyId=MyKMSKeyId
```

## カスタムキープロバイダーの作成

カスタムキープロバイダーの作成時に使用する暗号化のタイプに応じて、アプリケーションは異なる `EncryptionMaterialsProvider` インターフェイスも実装する必要があります。どちらのインターフェイスも AWS SDK for Java バージョン 1.11.0 以降で使用できます。

- Amazon S3 暗号化を実装するには、[com.amazonaws.services.s3.model.EncryptionMaterialsProvider interface](#) を使用します。

- ローカルディスク暗号化を実装するには、[com.amazonaws.services.elasticmapreduce.spi.security.EncryptionMaterialsProvider interface](https://docs.aws.amazon.com/emr/latest/ReleaseGuide/com.amazonaws.services.elasticmapreduce.spi.security.EncryptionMaterialsProvider.html) を使用します。

実装用の暗号化マテリアルを提供するには、任意の戦略を使用できます。例えば、静的暗号化マテリアルを提供するか、より複雑なキー管理システムと統合するかを選択できます。

Amazon S3 暗号化を使用している場合は、カスタム暗号化マテリアルに暗号化アルゴリズム AES/GCM/NoPadding を使用する必要があります。

ローカルディスク暗号化を使用している場合、カスタム暗号化マテリアルに使用する暗号化アルゴリズムは EMR リリースによって異なります。Amazon EMR 7.0.0 以前では、AES/GCM/NoPadding を使用する必要があります。Amazon EMR 7.1.0 以降では、AES を使用する必要があります。

EncryptionMaterialsProvider クラスは、暗号化コンテキストによって暗号化マテリアルを取得します。Amazon EMR は、呼び出し元が返す正しい暗号化マテリアルを判別しやすいように、実行時に暗号化コンテキスト情報を設定します。

Example 例: EMRFS での Amazon S3 の暗号化にカスタムキープロバイダを使用する

Amazon EMR が EncryptionMaterialsProvider クラスから暗号化マテリアルを取得して暗号化を実行すると、EMRFS はオプションで materialsDescription 引数にオブジェクトの Amazon S3 URI とクラスター JobFlowId の 2 つのフィールドを入力します。これは、EncryptionMaterialsProvider クラスが暗号化マテリアルを選択的に返すために使用できます。

たとえば、プロバイダは Amazon S3 URI プレフィックスごとに異なるキーを返すことができます。最終的に Amazon S3 オブジェクトに保存されるのは、EMRFS によって生成され、プロバイダに渡される materialsDescription 値ではなく、返された暗号化マテリアルの記述です。Amazon S3 オブジェクトの復号中に、暗号化マテリアルの説明が EncryptionMaterialsProvider クラスに渡されるため、再度、一致するキーを選択してオブジェクトを復号できます。

EncryptionMaterialsProvider リファレンス実装を以下に示します。別のカスタムプロバイダーである [EMRFSRSA EncryptionMaterialsProvider](#) は、から入手できます GitHub。

```
import com.amazonaws.services.s3.model.EncryptionMaterials;
import com.amazonaws.services.s3.model.EncryptionMaterialsProvider;
import com.amazonaws.services.s3.model.KMSEncryptionMaterials;
import org.apache.hadoop.conf.Configurable;
import org.apache.hadoop.conf.Configuration;

import java.util.Map;
```

```
/**
 * Provides KMSEncryptionMaterials according to Configuration
 */
public class MyEncryptionMaterialsProviders implements EncryptionMaterialsProvider,
Configurable{
    private Configuration conf;
    private String kmsKeyId;
    private EncryptionMaterials encryptionMaterials;

    private void init() {
        this.kmsKeyId = conf.get("my.kms.key.id");
        this.encryptionMaterials = new KMSEncryptionMaterials(kmsKeyId);
    }

    @Override
    public void setConf(Configuration conf) {
        this.conf = conf;
        init();
    }

    @Override
    public Configuration getConf() {
        return this.conf;
    }

    @Override
    public void refresh() {

    }

    @Override
    public EncryptionMaterials getEncryptionMaterials(Map<String, String>
materialsDescription) {
        return this.encryptionMaterials;
    }

    @Override
    public EncryptionMaterials getEncryptionMaterials() {
        return this.encryptionMaterials;
    }
}
```

## を使用したカスタムマテリアルプロバイダーの指定 AWS CLI

AWS CLIを使用するには、Encryption、ProviderType、CustomProviderClass、および CustomProviderLocation の各引数を emrfs オプションに渡します。

```
aws emr create-cluster --instance-type m5.xlarge --release-label emr-4.7.2 or earlier
--emrfs Encryption=ClientSide,ProviderType=Custom,CustomProviderLocation=s3://
mybucket/myfolder/provider.jar,CustomProviderClass=classname
```

Encryption を ClientSide に設定すると、クライアント側の暗号化が有効になります。CustomProviderClass は EncryptionMaterialsProvider オブジェクトの名前を表します。また、CustomProviderLocation は、ローカルまたは Amazon S3 のロケーションを示しています。この場所から Amazon EMR が CustomProviderClass をクラスター内の各ノードにコピーし、クラスパスに配置します。

## SDK を使用してカスタムマテリアルプロバイダーを指定する

SDK を使用するには、プロパティ fs.s3.cse.encryptionMaterialsProvider.uri を設定し、Amazon S3 に保存するカスタム EncryptionMaterialsProvider クラスをクラスター内の各ノードにダウンロードします。これは、emrfs-site.xml ファイルで設定します。このとき、CSE を有効にして、カスタムプロバイダーの適切な場所も指定します。

例えば、AWS SDK for Java を使用する では RunJobFlowRequest、コードは次のようになります。

```
<snip>
Map<String,String> emrfsProperties = new HashMap<String,String>();
    emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider.uri","s3://mybucket/
MyCustomEncryptionMaterialsProvider.jar");
    emrfsProperties.put("fs.s3.cse.enabled","true");
    emrfsProperties.put("fs.s3.consistent","true");

emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider","full.class.name.of.EncryptionMate

Configuration myEmrfsConfig = new Configuration()
    .withClassification("emrfs-site")
    .withProperties(emrfsProperties);

RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("Custom EncryptionMaterialsProvider")
    .withReleaseLabel("emr-7.1.0")
    .withApplications(myApp)
    .withConfigurations(myEmrfsConfig)
```

```

.withServiceRole("EMR_DefaultRole_V2")
.withJobFlowRole("EMR_EC2_DefaultRole")
.withLogUri("s3://myLogUri/")
.withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myEc2Key")
    .withInstanceCount(2)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m5.xlarge")
    .withSlaveInstanceType("m5.xlarge")
);

RunJobFlowResult result = emr.runJobFlow(request);
</snip>

```

## 引数を使用するカスタム EncryptionMaterialsProvider

プロバイダーに直接引数を渡す必要がある場合があります。これを行うには、プロパティとして定義されたカスタム引数を持つ `emrfs-site` 設定分類を使用できます。設定例を次に示します。これは、ファイル `myConfig.json` として保存されます。

```

[
  {
    "Classification": "emrfs-site",
    "Properties": {
      "myProvider.arg1": "value1",
      "myProvider.arg2": "value2"
    }
  }
]

```

から `create-cluster` コマンドを使用すると AWS CLI、次に示すように、`--configurations` オプションを使用して ファイルを指定できます。

```

aws emr create-cluster --release-label emr-7.1.0 --instance-type m5.xlarge
--instance-count 2 --configurations file://myConfig.json --emrfs
Encryption=ClientSide,CustomProviderLocation=s3://mybucket/myfolder/
myprovider.jar,CustomProviderClass=classname

```

## EMRFS S3EC V2 サポートを設定する

S3 Java SDK リリース (1.11.837 以降) では、さまざまなセキュリティ機能拡張がある暗号化クライアントバージョン 2 (S3EC V2) がサポートされます。詳細については、S3 のブログ投稿「[Updates](#)

[to the Amazon S3 encryption client](#)」を参照してください。また、「AWS SDK for Java デベロッパーガイド」の[Amazon S3 暗号化クライアントの移行](#)」を参照してください。

下位互換性を保つため、暗号化クライアント V1 は SDK で引き続き使用できます。デフォルトでは、CSE が有効になっている場合、EMRFS は S3EC V1 を使用して S3 オブジェクトを暗号化および復号します。

S3EC V2 で暗号化された S3 オブジェクトは、リリースバージョンが emr-5.31.0 より前 (emr-5.30.1 以前、emr-6.1.0 以前) の EMR クラスターでは、EMRFS によって復号化できません。

Example S3EC V2 を使用するように EMRFS を設定する

S3EC V2 を使用するように EMRFS を設定するには、次の設定を追加します。

```
{
  "Classification": "emrfs-site",
  "Properties": {
    "fs.s3.cse.encryptionV2.enabled": "true"
  }
}
```

## Amazon S3 クライアント側暗号化の `emrfs-site.xml` プロパティ

プロパティ	デフォルト値	説明
<code>fs.s3.cse.enabled</code>	<b>false</b>	<b>true</b> に設定すると、Amazon S3 に保存された EMRFS オブジェクトはクライアント側の暗号化を使用して暗号化されません。
<code>fs.s3.cse.encryptionV2.enabled</code>	<b>false</b>	<b>true</b> に設定すると、EMRFS は S3 暗号化クライアントバージョン 2 を使用して S3 上のオブジェクトの暗号化と復号化を行います。EMR バージョン 5.31.0 以降で使用できます。



プロパティ	デフォルト値	説明
<code>fs.s3.cse.encryptedMaterialsProvider.uri</code>	<b>N/A</b>	カスタム暗号化マテリアルを使用するときに適用されません。EncryptionMaterialsProvider を指定する JAR が保存されている Amazon S3 URI を意味します。この URI を指定すると、Amazon EMR により JAR がクラスター内のすべてのノードに自動的にダウンロードされます。
<code>fs.s3.cse.encryptedMaterialsProvider</code>	<b>N/A</b>	クライアント側の暗号化で使用する EncryptionMaterialsProvider のクラスパス。CSE-KMS を使用する場合は、 <code>com.amazon.ws.emr.hadoop.fs.cse.KMSEncryptionMaterialsProvider</code> を指定します。
<code>fs.s3.cse.materialsDescription.enabled</code>	<b>false</b>	に設定すると true、は暗号化されたオブジェクトの materialsDescription に、オブジェクトの Amazon S3 URI とを入力します JobFlowId。カスタム暗号化マテリアルを使用するときは true に設定します。
<code>fs.s3.cse.kms.keyId</code>	<b>N/A</b>	CSE-KMS を使用するとき適用されます。暗号化に使用される KMS キーの、KeyIdARN、またはエイリアスの値。

プロパティ	デフォルト値	説明
fs.s3.cse.cryptoStorageMode	<b>ObjectMetadata</b>	<p>Amazon S3 ストレージモード。デフォルトでは、暗号化情報の説明はオブジェクトメタデータに保存されます。インストラクションファイルに説明を保存することもできます。有効な値は ObjectMetadata と InstructionFile。詳細については、「<a href="#">AWS SDK for Java および Amazon S3 によるクライアント側のデータ暗号化</a>」を参照してください。</p>

# Amazon CloudWatch エージェント

Amazon EMR の Amazon CloudWatch エージェントは、EMR クラスター内の Amazon EC2 インスタンスをモニタリングできるツールです。エージェントで収集したメトリクスは、CloudWatch に保存して表示できます CloudWatch。CloudWatch エージェントの詳細については、[「Amazon ユーザーガイド CloudWatch」](#) を参照してください。

## Note

他の [Amazon CloudWatch エージェントメトリクス](#) AWS のサービスを公開、クエリ、または表示する場合、追加料金が発生します。料金の詳細については、以下のページを参照してください。

- [Amazon CloudWatch の料金](#)
- [Amazon Managed Service for Prometheus の料金](#)
- [Amazon Managed Grafana の料金](#)

これらのメトリクスは、Amazon EMR が AWS/ElasticMapReduce の名前空間で提供する無料のメトリクスとは別のものです CloudWatch。CloudWatch エージェントが公開しないメトリクスの詳細については、「[による Amazon EMR メトリクスのモニタリング CloudWatch](#)」を参照してください。

Amazon EMR 7.0 以降では、クラスターにカスタムバージョンの Amazon CloudWatch エージェントをインストールして、EMR クラスターからメトリクスを収集できます。エージェントは、クラスター内のすべてのノードからメトリクスを収集し、プライマリノードで収集して、そのノードからクラウドにメトリクスを発行します。

エージェントは、コンソールで新しいクラスターを作成するとき、または create-cluster API を使用するときインストールできます。詳細については、「[Amazon CloudWatch エージェントを使用する EMR クラスターを作成する](#)」を参照してください。デフォルトでは、リリース 7.x シリーズを実行する Amazon EMR クラスターは 60 秒間隔で [34 のシステムレベルのメトリクス](#) を CloudWatch に発行しますが、異なるメトリクスを発行するようにエージェントを設定できます。もう 1 つのオプションは、Amazon Managed Service for Prometheus にメトリクスを発行することであり、発行するメトリクスを選択することもできます。さまざまなユースケースやセットアップでは、CloudWatch コンソール、Amazon Managed Grafana、または API CloudWatch または Amazon

Managed Service for APIs を使用してメトリクスを表示およびクエリするようにエージェントを設定できます。メトリクスの保存とクエリに使用すると、料金 AWS のサービスが発生します。

Amazon EMR の CloudWatch エージェントは、Amazon EMR クラスター上の以下のサービスのシステムメトリクスと JMX メトリクスを発行できます。

- Hadoop DataNode
- Hadoop NameNode
- Yarn NodeManager
- ヤーン ResourceManager
- HBase Master – Amazon EMR 7.1 以降のみ
- HBase RegionServer – Amazon EMR 7.1 以降のみ
- HBase ThriftServer – Amazon EMR 7.1 以降のみ

使用可能なメトリクスと Amazon EMR で CloudWatch エージェントを設定する方法の詳細については、「[Amazon EMR の CloudWatch エージェントの設定](#)」を参照してください。

次の表に、Amazon EMR 7.x シリーズの最新リリース AmazonCloudWatchAgent に含まれているバージョンと、Amazon EMR がと共にインストールするコンポーネントを示します AmazonCloudWatchAgent。

このリリース AmazonCloudWatchAgent でと共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

#### AmazonCloudWatchAgent emr-7.1.0 のバージョン情報

Amazon EMR リリースラベル	AmazonCloudWatchAgent バージョン	でインストールされるコンポーネント AmazonCloudWatchAgent
emr-7.1.0	AmazonCloudWatchAgent 1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno

Amazon EMR リリースラベル	AmazonCloudWatchAgent バージョン	でインストールされるコンポーネント AmazonCloudWatchAgent  de, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
--------------------	--------------------------------	---

## トピック

- [Amazon CloudWatch エージェントを使用する EMR クラスターを作成する](#)
- [Amazon EMR を使用する CloudWatch エージェントのデフォルトメトリクス](#)
- [Amazon EMR の CloudWatch エージェントの設定](#)
- [考慮事項と制約事項](#)
- [CloudWatch エージェントリリース履歴](#)

## Amazon CloudWatch エージェントを使用する EMR クラスターを作成する

このセクションの手順では、AWS Management Console および から Amazon CloudWatch エージェントを使用して Amazon EMR でクラスターを作成する手順について説明します AWS CLI。

## トピック

- [CloudWatch エージェントに必要な IAM アクセス許可](#)
- [必要な CloudWatch エージェントエンドポイント](#)
- [EMR クラスターの作成](#)

## CloudWatch エージェントに必要な IAM アクセス許可

CloudWatch エージェントには、Amazon EMR の Amazon EC2 インスタンスプロファイルに AWS Identity and Access Management (IAM) アクセス `cloudwatch:PutMetricData` 許可が必要で

す。Amazon EMR のデフォルトロールにはすでにこのアクセス許可があります。デフォルトのロールは、AWS CLI を使用して から作成できます `aws emr create-default-roles`。詳細については、「Amazon EMR 管理ガイド」の「[クラスター EC2 インスタンスのサービスロール \(EC2 インスタンスプロファイル\)](#)」を参照してください。

以下の IAM ポリシー例には `cloudwatch:PutMetricData` アクセス許可が含まれています。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "cloudwatch:PutMetricData",
      "Resource": "*"
    }
  ]
}
```

## 必要な CloudWatch エージェントエンドポイント

プライベートサブネット内の EMR クラスター CloudWatch のメトリクスを に発行するには、CloudWatch エージェントエンドポイントを作成し、プライベートサブネットがある VPC に関連付けます。

各 の CloudWatch エンドポイントの詳細については AWS リージョン、「AWS 全般のリファレンスガイド」の「[Amazon CloudWatch エンドポイントとクォータ](#)」を参照してください。

## EMR クラスターの作成

CloudWatch エージェントで使用するために必要なアクセス許可とエンドポイントを設定したら、AWS Management Console または を使用して、エージェントがインストールされた新しいクラスター AWS CLI を作成します。

### Console

コンソールから Amazon CloudWatch エージェントを使用してクラスターを作成するには

1. Amazon EMR コンソールに移動します。
2. [クラスターを作成] を選択します。
3. [名前とアプリケーション] で、Amazon EMR リリース 7.0.0 以降を選択します。

4. アプリケーションバンドルで、クラスターにインストールするバンドルまたはアプリケーションを選択し、選択にCloudWatch エージェントを含めます。
5. ユースケースのニーズを満たすクラスターの作成に進みます。

## AWS CLI

では AWS CLI、の `--applications` パラメータを使用して Amazon CloudWatch エージェントをクラスターに追加できます `create-cluster`。

から Amazon CloudWatch エージェントを使用してクラスターを作成するには AWS CLI

- クラスターを作成するときは、次のようなコマンドを使用して Amazon CloudWatch エージェントを含めます。 `myKey` をお使いの EC2 キーペアの名前に置き換えます。

```
aws emr create-cluster --name "Spark cluster with CloudWatch agent" \  
--release-label emr-7.0.0 \  
--applications Name=Spark Name=AmazonCloudWatchAgent \  
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \  
--instance-count 3 --use-default-roles
```

で Amazon EMR を使用方法の詳細については AWS CLI、[AWS CLI 「コマンドリファレンス」](#) を参照してください。

## Amazon EMR を使用する CloudWatch エージェントのデフォルトメトリクス

Amazon EMR に Amazon CloudWatch エージェントをインストールすると、[エージェントを異なる方法で設定](#)しない限り、デフォルト設定によってクラスター内のすべてのインスタンスに対して次のシステムメトリクスが公開されます。各メトリクスの定義については、「[Amazon ユーザーガイド](#)」の [CloudWatch 「エージェントによって収集されるメトリクス」](#) を参照してください。 CloudWatch

### CPU

#### CPU メトリクス

- `cpu_usage_active`
- `cpu_usage_guest`

- `cpu_usage_guest_nice`
- `cpu_usage_idle`
- `cpu_usage_iowait`
- `cpu_usage_irq`
- `cpu_usage_nice`
- `cpu_usage_softirq`
- `cpu_usage_steal`
- `cpu_usage_system`
- `cpu_usage_user`

## Disk

### ディスクメトリクス

- `disk_free`
- `disk_total`
- `disk_used`
- `disk_used_percent`

## Memory

### メモリメトリクス

- `mem_active`
- `mem_available`
- `mem_available_percent`
- `mem_free`
- `mem_inactive`
- `mem_total`
- `mem_used`
- `mem_used_percent`
- `mem_buffered`



- `mem_cached`

## Network IO

### ネットワーク I/O メトリクス

- `net_bytes_recv`
- `net_bytes_sent`
- `net_packets_recv`
- `net_packets_sent`

## Process

### プロセスメトリクス

- `processes_running`
- `processes_total`

## Swap

### スワップのメトリクス

- `swap_free`
- `swap_used`
- `swap_used_percent`

デフォルトでは、エージェントはカスタム名前空間 CloudWatch の スキーマ、`CWAgent`、`instance.id`の `cluster.id`にすべてのメトリクスを発行します `node.typeservice.name`。エージェントが Amazon EMR CloudWatch に発行するこれらのデフォルトのメトリクスを表示するには、次のステップを実行します。

1. Amazon CloudWatch コンソールに移動します。
2. [メトリクス] タブを選択し、次に [すべてのメトリクス] を選択します。
3. [カスタム名前空間] で、**CWAgent** を選択します。次に、スキーマ **cluster.id**、**instance.id**、**node.typeservice.name**を選択します。
4. ユースケースに応じて、CloudWatch インターフェイスからメトリクスのクエリを続行します。

Amazon EMR は、各メトリクスに次のラベルをアタッチします: ["cluster.id", "instance.id", "node.type", "service.name"]。これらのラベルについては、次の点を考慮してください。

- cluster.id ラベルの値は、メトリクスを生成した EMR クラスターの ID です。例では、cluster.id ラベルの値は j-123456789ABC です。
- instance.id ラベルの値は、メトリクスを生成した EMR クラスター内のインスタンスの ID です。例では、instance.id ラベルの値は i-01bcf5f140f3355777 です。
- node.type ラベルの値は、メトリクスを生成したクラスター内のインスタンスのノードタイプを表します。この値は、master、core または task です。
- service.name ラベルの値は、メトリクスを生成したサービスの名前です。デフォルトシステムメトリックの service.name のデフォルト値は system です。

#### Note

Amazon EMR リリース 7.0.0 では、若干異なるメトリクスラベルスキーマが使用されます。cluster.id は jobflow.id、node.type は利用できません。instance.id と service.name は利用可能です。

## Amazon EMR の CloudWatch エージェントの設定

Amazon EMR 7.0.0 以降には、Amazon CloudWatch エージェントが含まれています。以下のセクションでは、「」の説明に従って EMR クラスターで エージェントを設定する方法について説明します [Amazon CloudWatch エージェントを使用する EMR クラスターを作成する](#)。

### トピック

- [Amazon EMR 7.1.0 の CloudWatch エージェントを設定する](#)
- [Amazon EMR 7.0.0 の CloudWatch エージェントを設定する](#)

## Amazon EMR 7.1.0 の CloudWatch エージェントを設定する

Amazon EMR 7.1.0 以降では、Amazon EMR 設定 API を使用して、追加のシステムメトリクスを使用するように Amazon CloudWatch エージェントを設定し、アプリケーションメトリクスを追加し、メトリクスの送信先を変更できます。EMR 設定 API を使用してクラスターのアプリケーションを設定する方法の詳細については、「[アプリケーションを設定する](#)」を参照してください。

**Note**

7.1.0 は、再設定タイプのみをサポートしますOVERWRITE。再設定タイプの詳細については、[「インスタンスグループを再設定する際の考慮事項」](#)を参照してください。

## トピック

- [設定スキーマ](#)
- [システムメトリクス設定の例](#)
- [アプリケーションメトリクス設定の例](#)
- [Amazon Managed Service for Prometheus の例](#)

## 設定スキーマ

emr-metrics には次の分類があります。

- `emr-system-metrics` — CPU、ディスク、メモリなどのシステムメトリクスを設定します。
- `emr-hadoop-hdfs-datanode-metrics` — Hadoop DataNode JMX メトリクスを設定する
- `emr-hadoop-hdfs-namenode-metrics` — Hadoop NameNode JMX メトリクスを設定する
- `emr-hadoop-yarn-nodemanager-metrics` — Yarn NodeManager JMX メトリクスを設定する
- `emr-hadoop-yarn-resourcemanager-metrics` — Yarn ResourceManager JMX メトリクスを設定する
- `emr-hbase-master-metrics` — HBase マスター JMX メトリクスを設定する
- `emr-hbase-region-server-metrics` — HBase リージョンサーバー JMX メトリクスの設定
- `emr-hbase-rest-server-metrics` — HBase REST Server JMX メトリクスの設定
- `emr-hbase-thrift-server-metrics` — HBase Thrift Server JMX メトリクスを設定する

次の表は、すべての分類で使用可能なプロパティと設定を示しています。

## emr-metrics プロパティ

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics_destination	オプションです。	クラスターメトリクスを Amazon CloudWatch に発行するか、Amazon Managed Service for Prometheus に発行するかを決定します。	「クラウドウォッチ」	「CLOUDWATCH」、「PROMETHEUS」	このプロパティでは、大文字と小文字は区別されません。例えば、「Cloudwatch」は「CLOUDWATCH」と同じです。
prometheus_endpoint	オプションです。	metrics_destination が「PROMETHEUS」に設定されている場合、このプロパティは、提供された Amazon Managed Service for Prometheus リモート書き込みエンドポイントにメトリクスを送信するように CloudWatch エージェント	該当なし	有効な Amazon Managed Service for Prometheus リモート書き込み URL。リモート書き込み URL 形式は <code>https://aps-workspaces.&lt;region&gt;.com/workspaces/&lt;workspace_id&gt;/api/v1/remote_write</code> です。	が「PROMETHEUS」に設定されている場合 metrics_destination、このフィールドは必須です。キーを指定しない場合、または値が空の文字列の場合、プロビジョニングは失敗します。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
		トを設定します。			

## emr-system-metrics プロパティ

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics_collection_interval	オプションです。	CloudWatch エージェントからメトリクスが収集および公開される秒単位の頻度。	「60」	秒数を指定する文字列。整数のみを使用できます。	このプロパティは、個々のメトリクスグループの metrics_collection_interval プロパティで上書きできます。

## emr-system-metrics 設定

### cpu

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics	オプションです。	エージェントが収集する CPU メトリクスのリスト。	「 <a href="#">Amazon EMR を使用した CloudWatch エージェントのデフォルトメトリクス</a> 」を参	usage_active や など、cpu_プレフィックスの有無にかかわらず、有効な CPU メトリクス名	空の文字列を指定すると、CPU メトリクスは発行されません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
			照してください。	のカンマ区切りリストcpu_time_idle 。有効な <a href="#">メトリクス</a> については、 <a href="#">CloudWatch「エージェントによって収集されるメトリクス」</a> を参照してください。	
metrics_collection_interval	オプションです。	エージェントが CPU メトリクスを収集して公開する秒単位の頻度。	グローバルの値metrics_collection_interval。	秒数を指定する文字列。整数のみを使用できます。	この値は、CPU メトリクスのグローバルmetrics_collection_interval プロパティのみを上書きします。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
drop_original_metrics	オプションです。	集計されていないメトリクスを発行しない CPU メトリクスのリスト。	未集計の CPU メトリクスは公開されていません。	メトリクスプロパティでも指定されている CPU メトリクスのカンマ区切りリスト。空の文字列は、すべての CPU メトリクスを発行することを意味します。	CloudWatch エージェントは、すべてのメトリクスをクラスター ID、インスタンス ID、ノードタイプ、サービス名別に集計します。デフォルトでは、CloudWatch エージェントは複数のリソースを持つメトリクスのリソースごとのメトリクスを公開しません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
resources	オプションです。	エージェントがコアごとのメトリクスを発行するかどうかを決定します。	"*"	「*」はコアごとのメトリクスを有効にします。「」はコアごとのメトリクスを無効にします。	CloudWatch エージェントは、で削除されない CPU メトリクスのコアごとのメトリクスのみを公開します drop_original_metrics 。

## disk

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics	オプションです。	エージェントが収集するディスクメトリクスのリスト。	<a href="#">「Amazon EMR を使用した CloudWatch エージェントのデフォルトメトリクス」</a> を参照してください。	disk_total やなどのdisk_プレフィックスの有無にかかわらず、有効なディスクメトリクス名のカンマ区切りリスト used_percent 。有効な <a href="#">メトリクスにつ</a>	空の文字列を指定すると、ディスクメトリクスが公開されません。



プロパティ	必要	説明	デフォルト値	使用できる値	メモ
				<p>いては、  <a href="#">CloudWatch</a>  <a href="#">「エージェントによっ</a>  <a href="#">て収集され</a>  <a href="#">るメトリク</a>  <a href="#">ス」</a>を参照            してください。</p>	
<code>metrics_collection_interval</code>	オプション です。	エージェントがディスクメトリクスを収集して公開する秒単位の頻度。	グローバルの値 <code>metrics_collection_interval</code> 。	秒数を指定する文字列。整数のみを使用できます。	この値は、ディスクメトリクスのグローバル <code>metrics_collection_interval</code> プロパティのみを上書きします。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
drop_original_metrics	オプションです。	集計されていないメトリクスを発行しないディスクメトリクスのリスト。	未集計のディスクメトリクスは公開されていません。	メトリクスプロパティでも指定されているディスクメトリクスのカンマ区切りリスト。空の文字列は、すべてのディスクメトリクスを発行することを意味します。	CloudWatch エージェントは、すべてのメトリクスをクラスター ID、インスタンス ID、ノードタイプ、サービス名別に集計します。デフォルトでは、CloudWatch エージェントは複数のリソースを持つメトリクスのリソースごとのメトリクスを公開しません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
resources	オプションです。	エージェントが per-mount-point メトリクスを発行するかどうかを決定します。	"*"	「*」はすべてのマウントポイントを意味し、「」はマウントポイントがないことを意味し、またはマウントポイントのカンマ区切りリストを意味します。例えば "/,/emr" です。	CloudWatch エージェントは、で削除されていないディスク per-mount-point メトリクスのメトリクスのみを公開します drop_original_metrics 。

## diskio

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics	オプションです。	エージェントが収集するディスク IO メトリクスのリスト。	<a href="#">「Amazon EMR を使用した CloudWatch エージェントのデフォルトメトリクス」</a> を参照してください。	diskio_reads やなどの diskio_プレフィックスの有無にかかわらず、有効なディスク IO メトリクス名のカンマ区切りリス	空の文字列を指定すると、ディスク IO メトリクスは発行されません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
				トwrites。 有効な <a href="#">メトリクス</a> については、 <a href="#">CloudWatch「エージェントによって収集されるメトリクス」</a> を参照してください。	
metrics_collection_interval	オプションです。	エージェントがディスク IO メトリクスを収集して発行する秒単位の頻度。	グローバルの値metrics_collection_interval。	秒数を指定する文字列。整数のみを使用できます。	この値は、ディスク IO メトリクスのグローバルmetrics_collection_interval プロパティのみを上書きします。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
drop_original_metrics	オプションです。	集計されていないメトリクスを発行しないディスク IO メトリクスのリスト。	未集約ディスク IO メトリクスは公開されていません。	メトリクスプロパティでも指定されているディスク IO メトリクスのカンマ区切りリスト。空の文字列は、すべてのディスク IO メトリクスを公開することを意味しません。	CloudWatch エージェントは、すべてのメトリクスをクラスター ID、インスタンス ID、ノードタイプ、サービス名別に集計します。デフォルトでは、CloudWatch エージェントは複数のリソースを持つメトリクスのリソースごとのメトリクスを公開しません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
resources	オプションです。	エージェントがデバイスごとのメトリクスを発行するかどうかを決定します。	"*"	「*」はすべてのストレージデバイスを意味し、「」はストレージデバイスがないこと、またはデバイス名のカンマ区切りリストを意味します。例えば "nvme0n1, nvme1n1" です。	CloudWatch エージェントは、で削除されないディスクIOメトリクスのデバイスごとのメトリクスのみを公開しますdrop_original_metrics。

## mem

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics	オプションです。	エージェントが収集するメモリメトリクスのリスト。	<a href="#">「Amazon EMR を使用した CloudWatch エージェントのデフォルトメトリクス」</a> を参照してください。	mem_available やなどのmem_プレフィックスの有無にかかわらず、有効なメモリメトリクス名のカンマ	空の文字列を指定すると、メモリメトリクスが公開されません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
				区切りリスト <code>available_percent</code> 。有効な <a href="#">メトリクスについては、CloudWatch「エージェントによって収集されるメトリクス」</a> を参照してください。	
<code>metrics_collection_interval</code>	オプションです。	エージェントがメモリメトリクスを収集して公開する秒単位の頻度。	グローバルの値 <code>metrics_collection_interval</code> 。	秒数を指定する文字列。整数のみを使用できます。	この値は、メモリメトリクスのグローバル <code>metrics_collection_interval</code> プロパティのみを上書きします。

## net

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
<code>metrics</code>	オプションです。	エージェントが収集するネット	<a href="#">「Amazon EMR を使用した</a>	<code>net_プレフィックス</code>	空の文字列を指定すると、ネット

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
		ワークメトリクスの一覧。	<a href="#">CloudWatch エージェントのデフォルトメトリクス</a> を参照してください。	の有無にかかわらず、有効なネットワークワークメトリクス名のカンマ区切りリスト。例: net_packets_sent および packets_received 。有効なメトリクスについては、 <a href="#">CloudWatch 「エージェントによって収集されるメトリクス」</a> を参照してください。	ワークメトリクスが公開されません。



プロパティ	必要	説明	デフォルト値	使用できる値	メモ
<code>metrics_collection_interval</code>	オプションです。	エージェントがネットワークメトリクスを収集して公開する秒単位の頻度。	グローバルの値 <code>metrics_collection_interval</code> 。	秒数を指定する文字列。整数のみを使用できます。	この値は、ネットワークメトリクスのグローバル <code>metrics_collection_interval</code> プロパティのみを上書きします。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
drop_original_metrics	オプションです。	集約されていないメトリクスを発行しないネットワークメトリクスのリスト。	未集約のネットワークメトリクスは公開されていません。	メトリクスプロパティでも指定されているネットワークメトリクスのカンマ区切りリスト。空の文字列は、すべてのネットワークメトリクスを公開することを意味します。	CloudWatch エージェントは、すべてのメトリクスをクラスター ID、インスタンス ID、ノードタイプ、サービス名別に集計します。デフォルトでは、CloudWatch エージェントは複数のリソースを持つメトリクスのリソースごとのメトリクスを公開しません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
resources	オプションです。	エージェントがインターフェイスごとのメトリクスを発行するかどうかを決定します。	"*"	「*」はすべてのネットワークインターフェイスを意味し、「」はネットワークインターフェイスがないこと、またはインターフェイス名のカンマ区切りリストを意味します。例えば "eth0,eth1" です。	CloudWatch エージェントは、で削除されないネットワークメトリクスのインターフェイスごとのメトリクスのみを公開します drop_original_metrics 。

## netstat

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics	オプションです。	エージェントが収集するネットワーク統計メトリクスのリスト。	<a href="#">「Amazon EMR を使用した CloudWatch エージェントのデフォルトメトリクス」</a> を参	tcp_listen や など、netstat_レフィックスの有無にかかわらず、有効なメモリ	空の文字列を指定すると、ネットワーク統計メトリクスが公開されません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
			照してください。	メトリクス名のカンマ区切りリストnetstat_udp_socket。有効なメトリクスについては、 <a href="#">CloudWatch「エージェントによって収集されるメトリクス」</a> を参照してください。	
metrics_collection_interval	オプションです。	エージェントがネットワーク統計メトリクスを収集して公開する秒単位の頻度。	グローバルの値metrics_collection_interval。	秒数を指定する文字列。整数のみを使用できます。	この値は、ネットワーク統計メトリクスのグローバルmetrics_collection_intervalプロパティのみを上書きします。

## processes

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics	オプションです。	エージェントが収集するプロセスメトリクスのリスト。	<a href="#">「Amazon EMR を使用した CloudWatch エージェントのデフォルトメトリクス」</a> を参照してください。	processes_running や など、processes_running フィックスの有無にかかわらず、有効なメモリメトリクス名のカンマ区切りリストtotal。有効なメトリクスについては、 <a href="#">CloudWatch 「エージェントによって収集されるメトリクス」</a> を参照してください。	空の文字列を指定すると、プロセスメトリクスが公開されません。
metrics_collection_interval	オプションです。	エージェントがシステムプロセスメトリクスを収集して公開する	グローバルの値metrics_collection_interval。	秒数を指定する文字列。整数のみを使用できます。	この値は、システムプロセスメトリクスのグローバルmetrics_c

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
		秒単位の頻度。			ollection_interval プロパティのみを上書きします。

## swap

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
metrics	オプションです。	エージェントが収集するスワップメトリクスのリスト。	<a href="#">「Amazon EMR を使用した CloudWatch エージェントのデフォルトメトリクス」</a> を参照してください。	swap_free や など、swap_プレフィックスの有無にかかわらず、有効なメモリメトリクス名のカンマ区切りリスト used_percent 。有効な <a href="#">メトリクスについては、CloudWatch 「エージェントによって収集されるメトリクス」</a> を参照	空の文字列を指定すると、スワップメトリクスは発行されません。

プロパティ	必要	説明	デフォルト値	使用できる値	メモ
				してください。	
metrics_collection_interval	オプションです。	エージェントがスワップメトリクスを収集して発行する秒単位の頻度。	グローバルの値 metrics_collection_interval。	秒数を指定する文字列。整数のみを使用できます。	この値は、スワップメトリクスのグローバル metrics_collection_interval プロパティのみを上書きします。

#### emr-hadoop-hdfs-datanodeメトリクスプロパティ

プロパティ	必要	説明	デフォルト値	使用できる値	
<code>&lt;custom_bean_name&gt;</code>	オプションです。	該当なし	CloudWatch エージェントがメトリクスを収集する MBean。 など Hadoop:service=DataNode,name=DataNodeActivity。 Amazon EMR リリース 7.0 の <a href="#">JMX YAML ファイルの</a>	MBean に関連付けられているメトリクスのカンマ区切りリストを含む文字列。例えば BlocksCached, BlocksRead です。	

プロパティ	必要	説明	デフォルト値	使用できる値
			<a href="#">例</a> で、サンプル MBean 名とそれに対応するメトリクスを確認できます。	
<code>otel.metric.export.interval</code>	オプションです。	Hadoop DataNode メトリクスを収集するミリ秒単位の頻度。	「60000」	ミリ秒数を指定する文字列。整数のみを使用できません。

#### emr-hadoop-hdfs-namenodeメトリクスプロパティ

プロパティ	必要	説明	デフォルト値	使用できる値
<code>&lt;custom_bean_name&gt;</code>	オプションです。	該当なし	CloudWatch エージェントがなどのメトリクスを収集する MBeanHadoop rvice=NameNode,nam e=FSNames ystem 。Amazon EMR リリース 7.0 の <a href="#">JMX YAML ファイルの例</a> で、サンプル MBean 名とそれに対応するメトリク	MBean に関連付けられているメトリクスのカンマ区切りリストを含む文字列。例えば BlockCapacity,CapacityUsedGB です。



プロパティ	必要	説明	デフォルト値	使用できる値
			スを確認できません。	
<code>otel.metric.export.interval</code>	オプションです。	Hadoop NameNode メトリクスを収集するミリ秒単位の頻度。	「60000」	ミリ秒数を指定する文字列。整数のみを使用できません。

### emr-hadoop-yarn-nodemanagerメトリクスプロパティ

プロパティ	必要	説明	デフォルト値	使用できる値
<code>&lt;custom_bean_name&gt;</code>	オプションです。	該当なし	CloudWatch エージェントがメトリクスを収集する MBean。などHadoop:service=NodeManager, name=NodeManagerMetrics。Amazon EMR リリース 7.0 の <a href="#">JMX YAML ファイルの例</a> で、サンプル MBean 名とそれに対応するメトリク	MBean に関連付けられているメトリクスのカンマ区切りリストを含む文字列。例えば MaxCapacity, AllocatedGB です。

プロパティ	必要	説明	デフォルト値	使用できる値
			スを確認できません。	
otel.metric.export.interval	オプションです。	Hadoop YARN NodeManager メトリクスを収集するミリ秒単位の頻度。	「60000」	ミリ秒数を指定する文字列。整数のみを使用できません。

### emr-hadoop-yarn-resourcemanagerメトリクスプロパティ

プロパティ	必要	説明	デフォルト値	使用できる値
<code>&lt;custom_bean_name&gt;</code>	オプションです。	該当なし	CloudWatch エージェントがメトリクスを収集する MBean。などHadoop:service=ResourceManager,name=PartitionQueueMetrics。Amazon EMR リリース 7.0 の <a href="#">JMX YAML ファイルの例</a> で、サンプル MBean 名とそれに対応	MBean に関連付けられているメトリクスのカンマ区切りリストを含む文字列。例えば MaxCapacity,MaxCapacityVCores です。

プロパティ	必要	説明	デフォルト値	使用できる値
			するメトリクスを確認できます。	
<code>otel.metric.export.interval</code>	オプションです。	Hadoop YARN ResourceManager メトリクスを収集するミリ秒単位の頻度。	「60000」	ミリ秒数を指定する文字列。整数のみを使用できます。

### emr-hbase-master-metrics プロパティ

プロパティ	必要	説明	デフォルト値	使用できる値
<code>&lt;custom_bean_name&gt;</code>	オプションです。	該当なし	CloudWatch エージェントがなどのメトリクスを収集する MBeanHadoop rvice=HBase, name=Master, sub=AssignmentManager。Amazon EMR リリース 7.0 の <a href="#">JMX YAML ファイルの例</a> で、サンプル MBean 名	MBean に関連付けられているメトリクスのカンマ区切りリストを含む文字列。例えば AssignFailedCount, AssignSubmittedCount です。

プロパティ	必要	説明	デフォルト値	使用できる値
			とそれに対応するメトリクスを確認できます。	
<code>otel.metric.export.interval</code>	オプションです。	HBase Master メトリクスを収集するミリ秒単位の頻度。	「60000」	ミリ秒数を指定する文字列。整数のみを使用できません。

### emr-hbase-region-serverメトリクスプロパティ

プロパティ	必要	説明	デフォルト値	使用できる値
<code>&lt;custom_mbean_name&gt;</code>	オプションです。	該当なし	CloudWatch エージェントがメトリクスを収集する MBean。などHadoop:service=HBase,name=RegionServer,sub=IPC。Amazon EMR リリース 7.0 の <a href="#">JMX YAML ファイルの例</a> で、サンプル MBean 名とそれに対応するメトリク	MBean に関連付けられているメトリクスのカンマ区切りリストを含む文字列。例えば <code>numActiveHandler,numActivePriorityHandler</code> です。

プロパティ	必要	説明	デフォルト値	使用できる値
			スを確認できません。	
<code>otel.metric.export.interval</code>	オプションです。	HBase リージョンサーバーのメトリクスを収集するミリ秒単位の頻度。	「60000」	ミリ秒数を指定する文字列。整数のみを使用できません。

### emr-hbase-rest-serverメトリクスプロパティ

プロパティ	必要	説明	デフォルト値	使用できる値
<code>&lt;custom_bean_name&gt;</code>	オプションです。	該当なし	CloudWatch エージェントがメトリクスを収集する MBean。などHadoop:service=HBase,name=REST。Amazon EMR リリース 7.0 の <a href="#">JMX YAML ファイルの例</a> で、サンプル MBean 名とそれに対応するメトリクスを確認できます。	MBean に関連付けられているメトリクスのカンマ区切りリストを含む文字列。例えば <code>successfulPut,successfulScanCount</code> です。

プロパティ	必要	説明	デフォルト値	使用できる値
<code>otel.metric.export.interval</code>	オプションです。	HBase Rest Server メトリクスを収集するミリ秒単位の頻度。	「60000」	ミリ秒数を指定する文字列。整数のみを使用できません。

### emr-hbase-thrift-serverメトリクスプロパティ

プロパティ	必要	説明	デフォルト値	使用できる値
<code>&lt;custom_mbean_name&gt;</code>	オプションです。	該当なし	CloudWatch エージェントがなどのメトリクスを収集する MBeanHadoop rvice=HBase, name=Thrift, sub=ThriftOne。Amazon EMR リリース 7.0 の <a href="#">JMX YAML ファイルの例</a> で、サンプル MBean 名とそれに対応するメトリクスを確認できます。	MBean に関連付けられているメトリクスのカンマ区切りリストを含む文字列。例えば BatchGet_max, BatchGet_mean です。
<code>otel.metric.export</code>	オプションです。	HBase Thrift サーバーメト	「60000」	ミリ秒数を指定する文字

プロパティ	必要	説明	デフォルト値	使用できる値
.interval		リクスを収集するミリ秒単位の頻度。		列。整数のみを使用できません。

## システムメトリクス設定の例

次の例は、すべてのシステムメトリクスのエクスポートを停止するように CloudWatch エージェントを設定する方法を示しています。

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

次の例では、デフォルトのシステムメトリクスをエクスポートするように CloudWatch エージェントを設定します。これにより、システムメトリクスを少なくとも 1 回再設定している場合に、エージェントをデフォルトのシステムメトリクスのみをエクスポートするようにリセットできます。このリセットにより、以前に再設定されたアプリケーションメトリクスも削除されます。

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": []
  }
]
```

次の例では、cpu、memおよび diskメトリクスをエクスポートするようにクラスターを設定します。

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {
          "metrics_collection_interval": "20"
        },
        "Configurations": [
          {
            "Classification": "cpu",
            "Properties": {
              "metrics": "cpu_usage_guest,cpu_usage_idle",
              "metrics_collection_interval": "30",
              "drop_original_metrics": "cpu_usage_guest"
            }
          },
          {
            "Classification": "mem",
            "Properties": {
              "metrics": "mem_active"
            }
          },
          {
            "Classification": "disk",
            "Properties": {
              "metrics": "disk_used_percent",
              "resources": "/,/mnt",
              "drop_original_metrics": ""
            }
          }
        ]
      }
    ]
  }
]
```

前の設定例には、次のプロパティがあります。



- 30 秒ごとに、エージェントはすべての CPU の `cpu_guest` メトリクスを収集します。CPU 集計メトリクスは CloudWatch、名前空間 にあります `CWAgent > cluster.id, instance.id, node.type, service.name`。
- 30 秒ごとに、エージェントはすべての CPU の `cpu_idle` メトリクスを収集します。CPU 集計メトリクスは CloudWatch、名前空間 にあります `CWAgent > cluster.id, instance.id, node.type, service.name`。エージェントは CPU ごとのメトリクスも収集します。これらは同じ名前空間にあります。 `drop_original_metrics` プロパティに が含まれていないため、エージェントはこのメトリクスを収集するため `cpu_idle`、エージェントはメトリクスを無視しません。
- 20 秒ごとに、エージェントは `mem_active` メトリクスを収集します。集計メトリクスは CloudWatch、名前空間 にあります `CWAgent > cluster.id, instance.id, node.type, service.name`。
- 20 秒ごとに、エージェントは `/` および `/mnt` ディスクマウントの `disk_used_percent` メトリクスを収集します。集約されたメトリクスは CloudWatch、名前空間 にあります `CWAgent > cluster.id, instance.id, node.type, service.name`。エージェントは、マウントごとのメトリクスも収集します。これらは同じ名前空間にあります。 `drop_original_metrics` プロパティに が含まれていないため、エージェントはこのメトリクスを収集するため `disk_used_percent`、エージェントはメトリクスを無視しません。

## アプリケーションメトリクス設定の例

次の例では、Hadoop Namenode サービスのメトリクスのエクスポートを停止するように CloudWatch エージェントを設定します。

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

次の例では、Hadoop アプリケーションメトリクスをエクスポートするようにクラスターを設定します。

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {
          "Hadoop:service=NameNode,name=FSNamesystem": "BlockCapacity,CapacityUsedGB",
          "otel.metric.export.interval": "20000"
        },
        "Configurations": []
      },
      {
        "Classification": "emr-hadoop-hdfs-datanode-metrics",
        "Properties": {
          "Hadoop:service=DataNode,name=JvmMetrics": "MemNonHeapUsedM",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      },
      {
        "Classification": "emr-hadoop-yarn-resourcemanager-metrics",
        "Properties": {
          "Hadoop:service=ResourceManager,name=CapacitySchedulerMetrics":
"AllocateNumOps,NodeUpdateNumOps"
        },
        "Configurations": []
      }
    ]
  }
]
```

前の例には、次のプロパティがあります。

- 20 秒ごとに、エージェントは Hadoop Namenode サービスを実行しているインスタンスから BlockCapacity および CapacityUsedGB メトリクスを収集します。
- 30 秒ごとに、エージェントは Hadoop Datanode サービスを実行しているインスタンスから MemNonHeapUsedM メトリクスを収集します。

- 30 秒ごとに、エージェントは Hadoop YARN を実行するインスタンスから AllocateNumOps および NodeUpdateNumOps メトリクスを収集します ResourceMaager。

## Amazon Managed Service for Prometheus の例

次の例は、Amazon Managed Service for Prometheus にメトリクスをエクスポートするように CloudWatch エージェントを設定する方法を示しています。

現在 Amazon Managed Service for Prometheus にメトリクスをエクスポートしていて、クラスターのメトリクスを再設定し、Amazon Managed Service for Prometheus へのメトリクスのエクスポートを続行する場合は、プロパティ `metrics_destination` と `prometheus_endpoint` を含める必要があります `prometheus_endpoint`。

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "http://amp-workspace/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

CloudWatch エージェントを使用してメトリクスを にエクスポートするには CloudWatch、次の例を使用します。

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
  }
]
```

**Note**

CloudWatch エージェントには、特定の属性の名前を変更する Prometheus エクスポートがあります。デフォルトのメトリクスラベルの場合、Amazon Managed Service for Prometheus は、Amazon が使用する期間の代わりにアンダースコア文字 CloudWatch を使用します。Amazon Managed Grafana を使用して Amazon Managed Service for Prometheus のデフォルトのメトリクスを視覚化すると、ラベルは `cluster_id`、`node_type`、および `instance_id` として表示されます `service_name`。

## Amazon EMR 7.0.0 の CloudWatch エージェントを設定する

デフォルトの CloudWatch エージェント設定で提供されるもの以外の追加のシステムメトリクスを使用するように Amazon エージェントを設定できます。[CloudWatch 7.0.0](#) の設定では、ブートストラップアクションを使用する必要があります。ブートストラップアクションについては、以下のセクションで例を挙げています。今後のリリースでは、Amazon EMR は Amazon EMR API を通じた追加の設定オプションを提供する予定です。

### トピック

- [Amazon EMR 7.0.0 を使用して追加のシステムメトリクスを設定する](#)
- [Amazon EMR 7.0.0 を使用してアプリケーションメトリクスを設定する](#)
- [Amazon EMR 7.0.0 でのメトリクスのクラウドストレージとして Amazon Managed Service for Prometheus を設定する](#)

## Amazon EMR 7.0.0 を使用して追加のシステムメトリクスを設定する

次のステップを使用して、Amazon EMR 7.0.0 の異なるシステムメトリクスセットを使用するようにエージェントを設定します。

1. CloudWatch エージェントメトリクスを指定する設定ファイルを保存する Amazon S3 アカウントでバケットを作成または選択します。
2. 希望するメトリクスを指定して `emr-amazon-cloudwatch-agent.json` 設定ファイルを作成します。これを行うには、[CloudWatch 「エージェント設定ファイルの作成」](#) で説明されている方法のいずれかを使用します。CloudWatch エージェント設定ファイルの構造の詳細については、「[Amazon CloudWatch ユーザーガイド](#)」の [CloudWatch 「エージェント設定ファイルを手動で作成または編集する」](#) を参照してください。

3. 次に、のaws-emr-utilitiesリポジトリに移動 GitHub し、次のシステムメトリクススクリプトをダウンロードします。
  - [install\\_system\\_metrics\\_launcher.sh](#) - ノードがブートストラップを完了できるように、install\_system\_metrics.sh をダウンロードしてバックグラウンドで実行するスクリプト。
  - [install\\_system\\_metrics.sh](#) - 実行中のインスタンスがブートストラップを完了するのを待ってから、JSON ファイル内の設定をダウンロードして適用するスクリプト。
4. 各 SH ファイルを開き、*my-s3-bucket* をステップ 1 のバケットの名前に置き換えます。
5. 1 つの JSON ファイルと 2 つの SH ファイルを S3 バケットにアップロードします。
6. これで、Amazon EMR コンソールに移動し、CloudWatch エージェントを使用して新しいクラスターを作成できます。左側のナビゲーションペインの [EMR on EC2] で [クラスター] を選択し、[クラスターを作成] を選択します。
7. [名前とアプリケーション] セクションで、Amazon EMR リリース 7.0.0 以降を選択します。
8. 「アプリケーションバンドル」で、クラスターにインストールするバンドルまたはアプリケーションを選択し、Amazon CloudWatch Agent を選択内容に含めます。
9. [ブートストラップアクション] セクションで [追加] を選択します。
  - [名前] には、**install\_system\_metrics\_launcher.sh** と入力します。
  - [スクリプトの場所] には、**s3://my-s3-bucket/**  
**install\_system\_metrics\_launcher.sh** を挿入します。*my-s3-bucket* をお使いの S3 バケットへのパス名に置き換えます。
  - [引数] ブロックは空のままにしておきます。
10. [ブートストラップアクションを追加] を選択します。
11. 引き続き、ワークロードのニーズに合わせてクラスターを作成してください。

クラスターが起動すると、CloudWatch エージェントは設定ファイルで指定したシステムメトリクスをに発行します CloudWatch。

## Amazon EMR 7.0.0 を使用してアプリケーションメトリクスを設定する

システムメトリクスに加えて、HDFS および YARN のアプリケーションメトリクスを発行するように Amazon CloudWatch エージェントを設定できます。アプリケーションメトリクスを公開するようにエージェントを設定するには、次の手順を実行します。

1. CloudWatch エージェントメトリクスを指定する設定ファイルを保存する Amazon S3 アカウントでバケットを作成または選択します。
2. 次に、の `aws-emr-utilities` リポジトリに移動 GitHub し、次のスクリプトをダウンロードします。
  - [install\\_app\\_metrics\\_launcher.sh](#) - ノードがブートストラップを完了できるように、`install_app_metrics.sh` をダウンロードしてバックグラウンドで実行するスクリプト。
  - [install\\_app\\_metrics.sh](#) - 実行中のインスタンスがブートストラップを完了するのを待ってから、次のステップでダウンロードする YAML ファイル内の設定をダウンロードして適用するスクリプト。
3. 各ファイルを開き、`my-s3-bucket` をステップ 1 のバケットの名前に置き換えます。
4. 次に、以下の YAML マッピングファイルをダウンロードします。これらの YAML ファイルの構造については、リポジトリ [javaagent](#) の OpenTelemetry Instrumentation for Java GitHub 「」を参照してください。
  - [datanode-metrics.yaml](#) - Hadoop DataNode メトリクスの設定。
  - [namenode-metrics.yaml](#) - Hadoop NameNode メトリクスの設定。
  - [nodemanager-metrics.yaml](#) - Yarn NodeManager メトリクスの設定。
  - [resourcemanager-metrics.yaml](#) - Yarn ResourceManager メトリクスの設定。
5. 2 つの SH ファイルと 4 つの YAML ファイルを S3 バケットにアップロードします。
6. これで、Amazon EMR コンソールに移動し、CloudWatch エージェントを使用して新しいクラスターを作成できます。左側のナビゲーションペインの [EMR on EC2] で [クラスター] を選択し、[クラスターを作成] を選択します。
7. [名前とアプリケーション] セクションで、Amazon EMR リリース 7.0.0 以降を選択します。
8. アプリケーションバンドル で、クラスターにインストールするアプリケーションのバンドルまたはカスタムグループを選択し、選択に CloudWatch エージェントを含めます。
9. [ブートストラップアクション] セクションで [追加] を選択します。
  - [名前] には、`install_app_metrics_launcher.sh` と入力します。
  - [スクリプトの場所] には、`s3://my-s3-bucket/install_app_metrics_launcher.sh` を挿入します。`my-s3-bucket` をお使いの S3 バケットへのパス名に置き換えます。
  - [引数] ブロックは空のままにしておきます。
10. [ブートストラップアクションを追加] を選択します。

11. 引き続き、ワークロードのニーズに合わせてクラスターを作成してください。

クラスターが起動すると、CloudWatch エージェントは指定したアプリケーションメトリクスとシステムメトリクスをに発行します CloudWatch。

## Amazon EMR 7.0.0 でのメトリクスのクラウドストレージとして Amazon Managed Service for Prometheus を設定する

の代わりに Amazon Managed Service for Prometheus にメトリクスを発行するように Amazon CloudWatch エージェントを設定できます CloudWatch。

### Note

Amazon CloudWatch エージェントから Amazon Managed Service for Prometheus または Amazon にメトリクスを発行することはできますが CloudWatch、同じクラスターの両方のサービスにメトリクスを発行することはできません。

Amazon Managed Service for Prometheus にメトリクスを発行するようにエージェントを設定するには、Amazon EMR の Amazon EC2 インスタンスプロファイルに `aps:RemoteWrite` AWS Identity and Access Management (IAM) アクセス許可を追加する必要があります。次のポリシー例は、必要なアクセス許可を含んでいます。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "aps:RemoteWrite",
      "Resource": "*"
    }
  ]
}
```

## EMR クラスターで CloudWatch エージェントを使用して Amazon Managed Service for Prometheus にメトリクスを発行する

サービスポリシーに正しいアクセス許可が付与されたら、次の手順を使用して、CloudWatch エージェントを使用してメトリクスを Amazon Managed Service for Prometheus に発行するクラスターを起動します。

1. AWS Management Console または AWS CLI を使用して、Amazon Managed Service for Prometheus ワークスペースを作成します。詳細については、「Amazon Managed Service for Prometheus ユーザーガイド」の「[ワークスペースの作成](#)」を参照してください。
2. Amazon S3 アカウントで、Amazon Managed Service for Prometheus をクラウドストレージとして指定する起動ファイルを保存するバケットを作成または選択します。
3. 次に、[aws-emr-utilities](#)リポジトリに移動し、次のスクリプトをダウンロードします。
  - [add\\_prometheus\\_endpoint\\_launcher.sh](#) - ノードがブートストラップを完了できるように、`add_prometheus_endpoint.sh` をダウンロードしてバックグラウンドで実行するスクリプト。
  - [add\\_prometheus\\_endpoint.sh](#) - 実行中のインスタンスがブートストラップを完了するのを待ってから、クラスターの起動時に引数として指定した Amazon Managed Service for Prometheus エンドポイントに発行するように CloudWatch エージェントを設定するスクリプト。
4. 各ファイルを開き、`my-s3-bucket` をステップ 2 のバケットの名前に置き換えます。
5. `aws emr create-cluster` を使用して AWS CLI、`add_prometheus_endpoint_launcher.sh` ブートストラップアクションで EMR クラスターを作成します。次のコマンドで、`my-s3-bucket` をブートストラップアクションを保持するバケットに置き換え、`managedpro-remote-write-workspace-url` を Amazon Managed Service for Prometheus ワークスペースのリモート書き込みエンドポイントに置き換えます。必ず `emr-7.0.0` またはそれ以降の Amazon EMR リリースラベルを指定してください。

```
aws emr create-cluster --name managedpro-cluster \  
  --release-label emr-7.0.0 \  
  --applications Name=Hadoop Name=AmazonCloudWatchAgent \  
  --ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \  
  --instance-count 3 --use-default-roles \  
  --bootstrap-actions Name='Add Prometheus Endpoint',Path=s3://my-s3-bucket/  
add_prometheus_endpoint_launcher.sh,Args='managedpro-remote-write-workspace-url'
```



クラスターが起動すると、CloudWatch エージェントは収集したメトリクスを Amazon Managed Service for Prometheus に発行します。

Amazon Managed Service for Prometheus を Amazon Managed Grafana 用データソースとして使用する

Amazon EMR が Amazon Managed Service for Prometheus にクラスターメトリクスを公開したら、次のステップを使用して Amazon Managed Grafana でメトリクスを視覚化できます。

1. を使用して AWS Management Console 、適切なアクセス許可を持つ Amazon Managed Grafana ワークスペースとユーザーを作成します。詳細については、「Amazon Managed Grafana ユーザーガイド」の「[ワークスペースの作成](#)」を参照してください。
2. Amazon Managed Service for Prometheus ワークスペースをデータソースとして Amazon Managed Grafana に追加します。詳細については、「Amazon Managed Grafana ユーザーガイド」の「[AWS データソース設定を使用して Amazon Managed Service for Prometheus をデータソースとして追加する](#)」を参照してください。

#### Note

CloudWatch エージェントには、特定の属性の名前を変更する Prometheus エクスポートがあります。デフォルトのメトリクスラベルの場合、Amazon Managed Service for Prometheus は、Amazon が使用する期間の代わりにアンダースコア文字 CloudWatch を使用します。そのため、Amazon Managed Grafana を使用して Amazon Managed Service for Prometheus のデフォルトメトリクスを視覚化すると、ラベルは `jobflow_id`、`instance_id`、`service_name` のように表示されます。

また、CloudWatch エージェントが Amazon Managed Service for Prometheus に発行するアプリケーションメトリクスは、`job`の代わりにラベルを使用します `service_name`。ただし、システムメトリクスは引き続き `service_name` のラベルを使用します。

## 考慮事項と制約事項

- ネイティブ Amazon CloudWatch エージェントは、Amazon EMR リリース 7.0.0 以降で作成したクラスターで使用できます。
- Amazon EMR 7.0.0 での CloudWatch エージェントの設定では、[ブートストラップアクション](#) を使用する必要があります。今後のリリースでは、Amazon EMR は Amazon EMR API を通じた追加の設定オプションを提供する予定です。

- クラスターを作成したリージョンに AWS Systems Manager エージェント (SSM CloudWatch エージェント) などの別の方法で CloudWatch エージェントをデプロイしている場合は、Amazon EMR エージェントをインストールできません。デプロイすると、クラスターはエラーで終了します。

## CloudWatch エージェントリリース履歴

次の表に、Amazon EMR の各リリースバージョン AmazonCloudWatchAgent に含まれる のバージョンと、アプリケーションと共にインストールされるコンポーネントを示します。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### AmazonCloudWatchAgent バージョン情報

Amazon EMR リリースラベル	AmazonCloudWatchAgent バージョン	と共にインストールされるコンポーネント AmazonCloudWatchAgent
emr-7.1.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.0.0	1.300031.1	adot-java-agent, emr-amazon-cloudwatch-agent

# Delta Lake

Delta Lake は、レイクハウスアーキテクチャのストレージレイヤーフレームワークであり、一般的に、Amazon S3 上に構築されます。Amazon EMR リリース 6.9.0 以降では、Delta Lake テーブルを使用する Amazon EMR クラスターで [Apache Spark 3.x](#) を使用できます。Delta Lake を使用するレイクハウスの詳細については、<https://delta.io/> を参照してください。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Delta のバージョンと、Amazon EMR で Delta と共にインストールされるコンポーネントを示しています。

このリリースで Delta と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Delta バージョン情報

Amazon EMR リリースラベル	Delta のバージョン	Delta と共にインストールされるコンポーネント
emr-7.1.0	Delta 3.0.0	Not available.

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Delta のバージョンと、Amazon EMR で Delta と共にインストールされるコンポーネントを示しています。

このリリースで Delta と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の Delta バージョン情報

Amazon EMR リリースラベル	Delta のバージョン	Delta と共にインストールされるコンポーネント
emr-6.15.0	Delta 2.4.0	Not available.

## Delta Lake の紹介

Delta Lake は、最新のデータレイクアーキテクチャを実装可能なオープンソースプロジェクトであり、一般的に、Amazon S3 上に構築されます。Delta Lake により、次の機能を得られます。

- Spark のトランザクション実行時に、原子性、一貫性、独立性、耐久性 (ACID) を確保できます。また、Spark ジョブの実行中に、一貫性のあるテーブル表示を行えます。
- Spark により、メタデータをスケーラブルに分散処理できます。
- ストリーミングとバッチのユースケースに、同じ Delta デルタテーブルで対応可能です。
- スキーマの自動適用によって、データインジェストの際に不良レコードを回避できます。
- データのバージョンングにより、特定の時点のデータを確認できます。
- 複雑なユースケースのマージ、更新、削除の操作 (例: 変更データキャプチャ (CDC)、ストリーミングの更新と挿入) に対応しています。

## Delta Lake がインストールされたクラスターを使用する

### トピック

- [Flink で Delta Lake クラスターを使用する](#)
- [Trino で Delta Lake クラスターを使用する](#)
- [Spark で Delta Lake クラスターを使用する](#)
- [Spark および AWS Glue と共に Delta Lake クラスターを使用する](#)

## Flink で Delta Lake クラスターを使用する

Amazon EMR リリース 6.11 以降では、Flink クラスターで Delta Lake を使用できます。次の例では AWS CLI、を使用して Amazon EMR Flink クラスターで Delta Lake を操作します。

### Note

Flink クラスターで Delta Lake を使用する場合、Amazon EMR は Flink DataStream API をサポートします。

## Delta Lake クラスターを作成する

1. 次のコンテンツを含む `delta_configurations.json` ファイルを作成します。

```
[{"Classification": "delta-defaults",  
  "Properties": {"delta.enabled": "true"}}]
```

2. 次の設定を使用して、クラスターを作成します。example Amazon S3 bucket path と subnet ID は、実際の値に置き換えてください。

```
aws emr create-cluster
--release-label emr-6.11.0
--applications Name=Flink
--configurations file://delta_configurations.json
--region us-east-1 --name My_Spark_Delta_Cluster
--log-uri s3://DOC-EXAMPLE-BUCKET/
--instance-type m5.xlarge
--instance-count 3
--service-role EMR_DefaultRole_V2
--ec2-attributes
InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Flink YARN セッションを初期化します。

Flink YARN セッションを初期化するには、次のコマンドを実行します。

```
flink-yarn-session -d
```

## Delta Lake を使用する Flink ジョブを作成する

次の例は、sbt または Maven を使用して Flink ジョブを作成し、Delta Lake を操作する方法を示しています。

sbt

[sbt](#) は Scala 用のビルドツールで、小規模なプロジェクトの場合、ほとんど、あるいは、まったく設定せずに使用できます。

```
libraryDependencies += Seq(
  "io.delta" %% "delta-flink" % deltaConnectorsVersion % "provided",
  "io.delta" %% "delta-standalone" % deltaConnectorsVersion % "provided",
  "org.apache.flink" %% "flink-clients" % flinkVersion % "provided",
  "org.apache.flink" %% "flink-parquet" % flinkVersion % "provided",
  "org.apache.hadoop" % "hadoop-client" % hadoopVersion % "provided",
  "org.apache.flink" % "flink-table-common" % flinkVersion % "provided",
  "org.apache.flink" %% "flink-table-runtime" % flinkVersion % "provided")
```

## Maven

[Maven](#) は、オープンソースのビルド自動化ツールで、Apache Software Foundation が提供しています。Maven では、Amazon EMR の Delta Lake を使用して、Flink ジョブを作成、公開、展開できます。

```
<project>
<properties>
  <scala.main.version>2.12</scala.main.version>
  <delta-connectors-version>0.6.0</delta-connectors-version>
  <flink-version>1.16.1</flink-version>
  <hadoop-version>3.1.0</hadoop-version>
</properties>

<dependencies>
  <dependency>
    <groupId>io.delta</groupId>
    <artifactId>delta-flink</artifactId>
    <version>${delta-connectors-version}</version>
    <scope>provided</scope>
  </dependency>
  <dependency>
    <groupId>io.delta</groupId>
    <artifactId>delta-standalone_${scala-main-version}</artifactId>
    <version>${delta-connectors-version}</version>
    <scope>provided</scope>
  </dependency>
  <dependency>
    <groupId>org.apache.flink</groupId>
    <artifactId>flink-clients</artifactId>
    <version>${flink-version}</version>
    <scope>provided</scope>
  </dependency>
  <dependency>
    <groupId>org.apache.flink</groupId>
    <artifactId>flink-parquet</artifactId>
    <version>${flink-version}</version>
    <scope>provided</scope>
  </dependency>
  <dependency>
    <groupId>org.apache.hadoop</groupId>
    <artifactId>hadoop-client</artifactId>
    <version>${hadoop-version}</version>
```

```

        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-common</artifactId>
        <version>${flink-version}</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-runtime</artifactId>
        <version>${flink-version}</version>
        <scope>provided</scope>
    </dependency>
</dependencies>

```

## Flink Datastream API を使用して、Delta テーブルに書き込む

次の例を使用して、を使用してテーブルに DeltaSink 書き込む を作成します。deltaTablePath:

```

public static DataStream<RowData> createDeltaSink(
    DataStream<RowData> stream,
    String deltaTablePath,
    RowType rowType) {
    Configuration configuration = new Configuration();
    DeltaSink<RowData> deltaSink = DeltaSink
        .forRowData(
            new org.apache.flink.core.fs.Path(deltaTablePath),
            configuration,
            rowType)
        .build();
    stream.sinkTo(deltaSink);
    return stream;
}

```

## Flink Datastream API を使用して、Delta テーブルから読み取る

次の例を使用して、を使用してテーブルから DeltaSource 読み取る 境界を作成します。deltaTablePath:

```

public static DataStream<RowData> createBoundedDeltaSourceAllColumns(

```

```
StreamExecutionEnvironment env,
String deltaTablePath) {
Configuration configuration = new Configuration();
DeltaSource<RowData> deltaSource = DeltaSource
    .forBoundedRowData(
        new org.apache.flink.core.fs.Path(deltaTablePath),
        configuration)
    .build();

return env.fromSource(deltaSource, WatermarkStrategy.noWatermarks(), "delta-
source");
}
```

## マルチクラスターサポートを使用してシンクを作成し、スタンドアロンの Delta Lake を操作する

次の例を使用して、`deltaTablePath`およびマルチクラスターサポート を使用してテーブルに DeltaSink 書き込む を作成します。 <https://docs.delta.io/latest/delta-standalone.html#multi-cluster-setup>

```
public DataStream<RowData> createDeltaSink(
    DataStream<RowData> stream,
    String deltaTablePath) {
    Configuration configuration = new Configuration();
    configuration.set("spark.delta.logStore.s3.impl",
        "io.delta.storage.S3DynamoDBLogStore");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.tableName",
        "delta_log");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.region", "us-
east-1");

    DeltaSink<RowData> deltaSink = DeltaSink
        .forRowData(
            new Path(deltaTablePath),
            configuration,
            rowType)
        .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```



## Flink ジョブを実行する

ジョブを実行するには、次のコマンドを実行します。

```
flink run FlinkJob.jar
```

## Trino で Delta Lake クラスターを使用する

Amazon EMR リリース 6.9.0 以降では、Trino クラスターで Delta Lake を使用できます。

このチュートリアルでは、を使用して Amazon EMR Trino クラスターの Delta Lake AWS CLI を操作します。

### Delta Lake クラスターを作成する

1. `delta_configurations.json` ファイルを作成し、選択したカタログの値を設定します。例えば、Hive メタストアをカタログとして使用する場合は、ファイルに次のコードを記述する必要があります。

```
[{"Classification":"delta-defaults",
  "Properties":{"delta.enabled":"true"}},
 {"Classification":"trino-connector-delta",
  "Properties":{"hive.metastore.uri":"thrift://localhost:9083"}}]
```

AWS Glue Catalog をストアとして使用する場合は、ファイルに次の内容が含まれている必要があります。

```
[{"Classification":"delta-defaults",
  "Properties":{"delta.enabled":"true"}},
 {"Classification":"trino-connector-delta",
  "Properties":{"hive.metastore":"glue"}}]
```

2. 次のように設定してクラスターを作成し、**example Amazon S3 bucket path** と **subnet ID** を実際の値に置き換えます。

```
aws emr create-cluster
  --release-label emr-6.9.0
  --applications Name=Trino
  --configurations file://delta_configurations.json
```

```
--region us-east-1 --name My_Spark_Delta_Cluster
--log-uri s3://DOC-EXAMPLE-BUCKET/
--instance-type m5.xlarge
--instance-count 2
--service-role EMR_DefaultRole_V2
--ec2-attributes
InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Delta Lake の Trino セッションを初期化します。

Trino セッションを初期化するには、次のコマンドを実行します。

```
trino-cli --catalog delta
```

## Delta Lake テーブルに書き込む

次の SQL コマンドを使用すると、テーブルを作成し、書き込みを行えます。

```
SHOW SCHEMAS;

CREATE TABLE default.delta_table (id int, data varchar, category varchar) WITH
(location = 's3://DOC-EXAMPLE-BUCKET/<prefix>');

INSERT INTO default.delta_table VALUES (1, 'a', 'c1'), (2, 'b', 'c2'), (3, 'c', 'c3');
```

## Delta Lake テーブルから読み取る

次の SQL コマンドを使用すると、テーブルから読み取りを行えます。

```
SELECT * from default.delta_table;
```

## Spark で Delta Lake クラスターを使用する

Amazon EMR バージョン 6.9.0 以降では、ブートストラップアクションを追加しなくても、Spark クラスターで Delta Lake を使用できます。Amazon EMR リリース 6.8.0 以前の場合、ブートストラップアクションを使用することで、依存関係上必要なファイルを事前にインストールできます。

次の例では AWS CLI、 を使用して Amazon EMR Spark クラスターで Delta Lake を操作します。

で Amazon EMR の Delta Lake を使用するには AWS Command Line Interface、まずクラスターを作成します。で Delta Lake 分類を指定する方法については AWS Command Line Interface、[「クラスターの作成 AWS Command Line Interface 時に 使用して設定を指定する」](#) または [「クラスターの作成時に Java SDK を使用して設定を指定する」](#) を参照してください。

1. 次のコンテンツを含む `configurations.json` ファイルを作成します。

```
[{"Classification":"delta-defaults", "Properties":{"delta.enabled":"true"} ]]
```

2. 次のように設定してクラスターを作成し、サンプルの Amazon S3 **bucket path** と **subnet ID** を実際の値に置き換えます。

```
aws emr create-cluster
  --release-label emr-6.9.0
  --applications Name=Spark
  --configurations file://delta_configurations.json
  --region us-east-1
  --name My_Spark_Delta_Cluster
  --log-uri s3://DOC-EXAMPLE-BUCKET/
  --instance-type m5.xlarge
  --instance-count 2
  --service-role EMR_DefaultRole_V2
  --ec2-attributes
  InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

次のファイルを、Spark ジョブに依存関係上必要な jar として使用することで、Amazon EMR クラスターと Spark アプリケーションを構築できます。

```
/usr/share/aws/delta/lib/delta-core.jar,
/usr/share/aws/delta/lib/delta-storage.jar,
/usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar
```

#### Note

Amazon EMR リリース 6.9.0 以降を使用する場合は、`/usr/share/aws/delta/lib/delta-spark.jar` の代わりに `/usr/share/aws/delta/lib/delta-core.jar` を使用します。

詳細については、「[Submitting Applications](#)」を参照してください。

Spark ジョブに依存関係上必要な jar を指定するには、次の設定プロパティを Spark アプリケーションに追加します。

```
--conf "spark.jars=/usr/share/aws/delta/lib/delta-core.jar,  
/usr/share/aws/delta/lib/delta-storage.jar,  
/usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Spark ジョブの依存関係については、「[Dependency Management](#)」を参照してください。

Amazon EMR リリース 6.9.0 以降を使用する場合は、代わりに `/usr/share/aws/delta/lib/delta-spark.jar` 設定を追加します。

```
--conf "spark.jars=/usr/share/aws/delta/lib/delta-spark.jar,  
/usr/share/aws/delta/lib/delta-storage.jar,  
/usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

## Delta Lake の Spark セッションを初期化する

次の例は、インタラクティブな Spark シェルを起動する方法、Spark 送信を使用する方法、Amazon EMR Notebooks を使用して Amazon EMR の Delta Lake を操作する方法をそれぞれ示しています。

spark-shell

1. SSH を使用してプライマリノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してプライマリノードに接続する](#)」を参照してください。
2. 以下のコマンドを入力して、Spark シェルを起動します。PySpark シェルを使用するには、`spark-shell` に置き換えます `pyspark`。

```
spark-shell \  
  --conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension" \  
  --conf  
  "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

## spark-submit

1. SSH を使用してプライマリノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してプライマリノードに接続する](#)」を参照してください。
2. Delta Lake の Spark セッションを起動するには、次のコマンドを入力します。

```
spark-submit
  -conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension"
  -conf
"spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

## EMR Studio notebooks

Amazon EMR Studio ノートブックを使用して Spark セッションを初期化するには、次の例のように、Amazon EMR Notebooks で %%configure マジックコマンドを使用して Spark セッションを設定します。詳細については、「Amazon EMR 管理ガイド」の「[EMR Notebooks マジックを使用する](#)」を参照してください。

```
%%configure -f
{
  "conf": {
    "spark.sql.extensions": "io.delta.sql.DeltaSparkSessionExtension",
    "spark.sql.catalog.spark_catalog":
"org.apache.spark.sql.delta.catalog.DeltaCatalog"
  }
}
```

## Delta Lake テーブルに書き込む

次の例は、を作成して Delta Lake データセットとして DataFrame 書き込む方法を示しています。また、デフォルトの Hadoop ユーザーとして、プライマリノードに SSH 接続し、Spark シェルでデータセットを操作する方法を示しています。

### Note

コードサンプルを Spark シェルに貼り付けるには、プロンプトで「:paste」と入力し、例を貼り付けて、[CTRL + D] を押します。

## PySpark

Spark には、Python ベースのシェルである pyspark も用意されており、Python で記述した Spark プログラムのプロトタイプ作成に使用できます。spark-shell の場合と同様に、プライマリノードで pyspark を呼び出します。

```
## Create a DataFrame
data = spark.createDataFrame([("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
                              ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
                              ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
                              ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")],
                              ["id", "creation_date", "last_update_time"])

## Write a DataFrame as a Delta Lake dataset to the S3 location
spark.sql("""CREATE TABLE IF NOT EXISTS delta_table (id string, creation_date
string,
last_update_time string)
USING delta location
's3://DOC-EXAMPLE-BUCKET/example-prefix/db/delta_table'""");

data.writeTo("delta_table").append()
```

## Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame
val data = Seq(("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
              ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
              ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
              ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")).toDF("id", "creation_date",
"last_update_time")

// Write a DataFrame as a Delta Lake dataset to the S3 location
spark.sql("""CREATE TABLE IF NOT EXISTS delta_table (id string,
creation_date string,
last_update_time string)
USING delta location
's3://DOC-EXAMPLE-BUCKET/example-prefix/db/delta_table'""");

data.write.format("delta").mode("append").saveAsTable("delta_table")
```

## SQL

```
-- Create a Delta Lake table with the S3 location
CREATE TABLE delta_table(id string,
creation_date string,
last_update_time string)
USING delta LOCATION
's3://DOC-EXAMPLE-BUCKET/example-prefix/db/delta_table';

-- insert data into the table
INSERT INTO delta_table VALUES ("100", "2015-01-01",
"2015-01-01T13:51:39.340396Z"),
("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103", "2015-01-01", "2015-01-01T13:51:40.519832Z");
```

## Delta Lake テーブルから読み取る

### PySpark

```
ddf = spark.table("delta_table")
ddf.show()
```

### Scala

```
val ddf = spark.table("delta_table")
ddf.show()
```

## SQL

```
SELECT * FROM delta_table;
```

## Spark および AWS Glue と共に Delta Lake クラスターを使用する

AWS Glue Catalog を Delta Lake テーブルのメタストアとして使用するには、次の手順でクラスターを作成します。を使用して Delta Lake 分類を指定する方法については [AWS Command Line Interface](#)、[「クラスターの作成 AWS Command Line Interface 時に を使用して設定を指定する」](#) または [「クラスターの作成時に Java SDK を使用して設定を指定する」](#) を参照してください。

## Delta Lake クラスターを作成する

1. 次のコンテンツを含む `configurations.json` ファイルを作成します。

```
[{"Classification":"delta-defaults",
"Properties":{"delta.enabled":"true"}},
{"Classification":"spark-hive-site",
"Properties":
{"hive.metastore.client.factory.class":"com.amazonaws.glue.catalog.metastore.AWSGlueDataCat
```

2. 次のように設定してクラスターを作成し、**example Amazon S3 bucket path** と **subnet ID** を実際の値に置き換えます。

```
aws emr create-cluster
  --release-label emr-6.9.0
  --applications Name=Spark
  --configurations file://delta_configurations.json
  --region us-east-1
  --name My_Spark_Delta_Cluster
  --log-uri s3://DOC-EXAMPLE-BUCKET/
  --instance-type m5.xlarge
  --instance-count 2
  --service-role EMR_DefaultRole_V2
  --ec2-attributes
  InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

## 考慮事項と制約事項

- Amazon EMR リリース 6.9.0 以降の場合、Delta Lake の使用がサポートされています。Amazon EMR クラスターでは、[Apache Spark 3.x](#) と Delta テーブルを使用できます。
- パフォーマンス、セキュリティ、信頼性を最適化するには、S3 ロケーションパスに `s3a` ではなく `s3` URI スキームを使用すると良いでしょう。詳細については、「[ストレージシステムとファイルシステムで作業する](#)」を参照してください。
- Amazon EMR 7.0 では、Delta Universal Format (UniForm) および `convert-to-Iceberg` ステートメントはサポートされていません。
- Amazon EMR 6.9 6.10 の場合、Amazon S3 に保存している Delta Lake のテーブルデータで列の名前を変更すると、列のデータが NULL になります。Amazon EMR 6.11 以降、この問題は解決



されています。列名変更の実験的な操作の詳細については、「[Delta Lake User Guide](#)」にある「Rename columns」の操作を参照してください。

- Apache Spark の外部で AWS Glue データカタログにデータベースを作成すると、データベースに空の LOCATION フィールドが含まれる可能性があります。Spark では空のロケーションプロパティを使用してデータベースを作成することはできません。そのため、Amazon EMR の Spark で Glue データベースに Delta テーブルが作成された後に、データベースに空の LOCATION プロパティが存在すると、次のエラーが発生します。

```
IllegalArgumentException: Can not create a Path from an empty string
```

この問題を解決するには、LOCATION フィールドに空ではない有効なパスを指定して、Data Catalog にデータベースを作成します。このソリューションを実装する手順については、「Amazon Athena ユーザーガイド」の「[テーブル作成時の不正な引数の例外](#)」を参照してください。

## Delta リリース履歴

次の表は、Amazon EMR の各リリースバージョンに含まれている Delta のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Delta バージョン情報

Amazon EMR リリースラベル	Delta のバージョン	Delta と共にインストールされるコンポーネント
emr-7.1.0	3.0.0	Not available.
emr-7.0.0	3.0.0	Not available.
emr-6.15.0	2.4.0	Not available.
emr-6.14.0	2.4.0	Not available.
emr-6.13.0	2.4.0	Not available.

Amazon EMR リリースラベル	Delta のバージョン	Delta と共にインストールされるコンポーネント
emr-6.12.0	2.4.0	Not available.
emr-6.11.1	2.2.0	Not available.
emr-6.11.0	2.2.0	Not available.
emr-6.10.1	2.2.0	Not available.
emr-6.10.0	2.2.0	Not available.
emr-6.9.1	2.1.0	Not available.
emr-6.9.0	2.1.0	Not available.

# Apache Flink

[Apache Flink](#) はストリーミングデータフローエンジンで、これを使用して、高スループットのデータソースでリアルタイムのストリーム処理ができます。Flink は、イベント、厳密に 1 回限りのセマンティクス、バックプレッシャー制御、ストリーミングアプリケーションとバッチアプリケーションの両方を書き込むように最適化された APIs out-of-order イベント時間セマンティクスをサポートしています。

さらに Flink には、次のようなサードパーティーのデータソースへのコネクタがあります。

- [Amazon Kinesis Data Streams](#)
- [Apache Kafka](#)
- [Flink Elasticsearch コネクタ](#)
- [Twitter ストリーミング API](#)
- [Cassandra](#)

Amazon EMR では、クラスター内の他のアプリケーションと一緒にリソースを管理できるように、YARN アプリケーションとして Flink をサポートしています。Flink を搭載した YARN では、一時的な Flink のジョブを送信できます。あるいは、全体的な YARN 予約に従って複数のジョブを受け入れ、リソースを割り当てる長期実行クラスターを作成することもできます。

Flink は、Amazon EMR リリースバージョン 5.1.0 以降に含まれています。

## Note

Amazon EMR リリースバージョン 5.2.1 では `FlinkKinesisConsumer` クラスのサポートが追加されました。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Flink のバージョンと、Amazon EMR で Flink と共にインストールされるコンポーネントを示しています。

このリリースで Flink と共にインストールされるコンポーネントのバージョンについては、[「リリース 7.1.0 コンポーネントバージョン」](#)を参照してください。

## emr-7.1.0 の Flink バージョン情報

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-7.1.0	Flink 1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Flink のバージョンと、Amazon EMR で Flink と共にインストールされるコンポーネントを示しています。

このリリースで Flink と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の Flink バージョン情報

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-6.15.0	Flink 1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client,

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
		flink-jobmanager-config, hudi, delta-standalone-connectors

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Flink のバージョンと、Amazon EMR で Flink と共にインストールされるコンポーネントを示しています。

このリリースで Flink と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

#### emr-5.36.2 の Flink バージョン情報

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.36.2	Flink 1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

#### トピック

- [Flink を使用してクラスターを作成する](#)
- [Amazon EMR で Flink を設定する](#)
- [Amazon EMR 内で Flink ジョブを操作する](#)
- [Scala シェルを使用する](#)
- [Flink ウェブインターフェイスを検索する](#)
- [Flink オートスケーラー](#)

- [タスク回復とスケーリング操作のためのジョブ再起動時間の最適化](#)
- [Amazon EMR の Zeppelin から Flink ジョブを操作する](#)
- [Flink リリース履歴](#)

## Flink を使用してクラスターを作成する

クラスターは、AWS Management Console、AWS CLI、または AWS SDK を使用して起動できます。

Flink がインストールされたクラスターをコンソールから起動するには

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Software Configuration (ソフトウェア設定)] で、EMR Release emr-5.1.0 以降を選択します。
4. インストールする他のアプリケーションとともに、アプリケーションに [Flink] を選択します。
5. 必要に応じて他のオプションを選択し、[Create cluster (クラスターの作成)] を選択します。

から Flink を使用してクラスターを起動するには AWS CLI

- 次のコマンドを使用してクラスターを作成します。

```
aws emr create-cluster --release-label emr-7.1.0 \  
--applications Name=Flink \  
--configurations file:///./configurations.json \  
--region us-east-1 \  
--log-uri s3://myLogUri \  
--instance-type m5.xlarge \  
--instance-count 2 \  
--service-role EMR_DefaultRole_V2 \  
--ec2-attributes KeyName=MyKeyName,InstanceProfile=EMR_EC2_DefaultRole \  
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\  
Args=flink-yarn-session,-d
```

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

## Amazon EMR で Flink を設定する

### Hive Metastore と Glue Catalog を使用して Flink を設定する

Amazon EMR リリース 6.9.0 以降では、Hive への Apache Flink コネクタを備えた Hive メタストアと AWS Glue Catalog の両方がサポートされています。このセクションでは、[AWS Glue Catalog](#) と [Hive Metastore](#) を Flink で使用するために必要な手順について概説します。

#### トピック

- [Hive Metastore を使用する](#)
- [AWS Glue データカタログを使用する](#)

#### Hive Metastore を使用する

1. リリース 6.9.0 以降の EMR でクラスターを作成し、少なくとも 2 つのアプリケーション (Hive と Flink) を使用します。
2. [スクリプトランナー](#)を使用して、次のスクリプトをステップ関数として実行します。

```
hive-metastore-setup.sh
```

```
sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
```

**Add step** ✕

**Step type** Custom JAR

**Name\***

**JAR location\***  JAR location maybe a path into S3 or a fully qualified java class in the classpath.

**Arguments**  These are passed to the main function in the JAR. If the JAR does not specify a main class in its manifest file you can specify another class name as the first argument.

**Action on failure**  What happens if the step fails

Cancel
Save

## AWS Glue データカタログを使用する

1. リリース 6.9.0 以降の EMR でクラスターを作成し、少なくとも 2 つのアプリケーション (Hive と Flink) を使用します。
2. AWS Glue Data Catalog の設定にある [Hive テーブルメタデータに使用] を選択して、クラスターでデータカタログを有効にします。
3. [スクリプトランナー](#)を使用して、次のスクリプトをステップ関数として実行します (「[Amazon EMR クラスターでのコマンドとスクリプトの実行](#)」を参照)。

glue-catalog-setup.sh

```

sudo cp /usr/lib/hive/auxlib/aws-glue-datacatalog-hive3-client.jar /usr/lib/flink/lib
sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
sudo chmod 755 /usr/lib/flink/lib/aws-glue-datacatalog-hive3-client.jar
sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar

```



**Add step**
✕

**Step type** Custom JAR

**Name\***

**JAR location\***  JAR location maybe a path into S3 or a fully qualified java class in the classpath.

**Arguments**  These are passed to the main function in the JAR. If the JAR does not specify a main class in its manifest file you can specify another class name as the first argument.

**Action on failure**  What happens if the step fails

Cancel
Save

## 設定ファイルを使用して Flink を構成する

Amazon EMR 設定 API と設定ファイルを使用して、Flink を構成できます。API 内で設定可能なファイルは次のとおりです。

- flink-conf.yaml
- log4j.properties
- flink-log4j-session
- log4j-cli.properties

flink-conf.yaml は、Flink の主要な設定ファイルです。

Flink に使用するタスクスロットの数を AWS CLI から設定するには

1. 次のコンテンツを含む configurations.json ファイルを作成します。

```
[
  {
    "Classification": "flink-conf",
    "Properties": {
      "taskmanager.numberOfTaskSlots": "2"
    }
  }
]
```

]

2. 次に、次の設定でクラスターを作成します。

```
aws emr create-cluster --release-label emr-7.1.0 \  
--applications Name=Flink \  
--configurations file://./configurations.json \  
--region us-east-1 \  
--log-uri s3://myLogUri \  
--instance-type m5.xlarge \  
--instance-count 2 \  
--service-role EMR_DefaultRole_V2 \  
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

### Note

Flink API でも、一部の設定を変更できます。詳細については、Flink のドキュメントで「[Concepts \(概念\)](#)」を確認してください。

Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## Parallelism オプション

アプリケーションの所有者であれば、Flink 内のタスクにどのリソースを割り当てるべきかを熟知しているでしょう。このドキュメントの例では、アプリケーションに使用するスレーブインスタンスと同じ数のタスクを使用します。初期レベルの並列処理では、多くの場合、そうすることをお勧めしますが、タスクスロットを使用して並列処理の粒度を高めることも可能です。ただし、一般的には、並列処理がインスタンスあたりの[仮想コア](#)数 (性能) を超えないようにする必要があります。Flink アーキテクチャの詳細については、Flink ドキュメントの「[Concepts](#)」を参照してください。

## 複数のプライマリノードを持つ EMR クラスターで Flink を設定する

Flink JobManager のは、複数のプライマリノードを持つ Amazon EMR クラスターのプライマリノードフェイルオーバープロセス中も引き続き使用できます。Amazon EMR 5.28.0 以降では、JobManager 高可用性も自動的に有効になっています。手動設定は必要ありません。

Amazon EMR バージョン 5.27.0 以前では、JobManager は単一障害点です。が JobManager 失敗すると、すべてのジョブ状態が失われ、実行中のジョブは再開されません。次の例に示すように、アプリケーションの試行回数、チェックポイント、Flink の状態ストレージ ZooKeeper として を有効にすることで、JobManager 高可用性を有効にできます。

```
[
  {
    "Classification": "yarn-site",
    "Properties": {
      "yarn.resourcemanager.am.max-attempts": "10"
    }
  },
  {
    "Classification": "flink-conf",
    "Properties": {
      "yarn.application-attempts": "10",
      "high-availability": "zookeeper",
      "high-availability.zookeeper.quorum": "%{hiera('hadoop:zk')}",
      "high-availability.storageDir": "hdfs:///user/flink/recovery",
      "high-availability.zookeeper.path.root": "/flink"
    }
  }
]
```

YARN の最大アプリケーションマスター試行回数と Flink のアプリケーション試行回数の両方を設定する必要があります。詳細については、「[YARN クラスターの高可用性の設定](#)」を参照してください。また、以前に完了したチェックポイントから JobManager 実行中のジョブを再開するように Flink チェックポイントを設定することもできます。詳細については、「[Flink のチェックポイント](#)」を参照してください。

## メモリプロセスサイズの設定

Flink 1.11.x を使用する Amazon EMR バージョンの場合、`jobmanager.memory.process.size` と TaskManager (`taskmanager.memory.process.size`) の両方 JobManager の合計メモリプロセスサイズを設定する必要があります `flink-conf.yaml`。これらの値を設定するには、設定 API を使用してクラスターを設定するか、SSH を介してこれらのフィールドのコメントを手動で解除します。Flink には、次のデフォルト値が設定されています。

- `jobmanager.memory.process.size: 1600m`

- `taskmanager.memory.process.size`: 1728m

JVM メタスペースとオーバーヘッドを除外する場合は、`taskmanager.memory.process.size` の代わりに Flink の合計メモリサイズ (`taskmanager.memory.flink.size`) を使用します。`taskmanager.memory.process.size` のデフォルト値は 1280m です。`taskmanager.memory.process.size` と `taskmanager.memory.process.size` の両方を設定することはお勧めしません。

Flink 1.12.0 以降を使用するいずれの Amazon EMR バージョンにも、Flink のオープンソースにリストされているデフォルト値が Amazon EMR のデフォルト値として設定されているため、自分で設定する必要はありません。

## ログ出力ファイルサイズの設定

Flink アプリケーションコンテナは、`.out` ファイル、`.log` ファイル、`.err` ファイルの 3 種類のログファイルを作成して書き込みます。`.err` ファイルのみが圧縮されてファイルシステムから削除され、`.log` および `.out` ログファイルはファイルシステムに残ります。これらの出力ファイルを管理しやすく、クラスターを安定させるために、`log4j.properties` のログローテーションを設定して、ファイルの最大数を設定し、それらのサイズを制限できます。

Amazon EMR バージョン 5.30.0 以降

Amazon EMR 5.30.0 以降の場合、Flink では、設定分類名 `flink-log4j` を持つ `log4j2` ロギングフレームワークが使用されます。次の設定例は、`log4j2` 形式を示しています。

```
[
  {
    "Classification": "flink-log4j",
    "Properties": {
      "appender.main.name": "MainAppender",
      "appender.main.type": "RollingFile",
      "appender.main.append" : "false",
      "appender.main.fileName" : "${sys:log.file}",
      "appender.main.filePattern" : "${sys:log.file}.%i",
      "appender.main.layout.type" : "PatternLayout",
      "appender.main.layout.pattern" : "%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p %-60c %x - %m%n",
      "appender.main.policies.type" : "Policies",
      "appender.main.policies.size.type" : "SizeBasedTriggeringPolicy",
      "appender.main.policies.size.size" : "100MB",
      "appender.main.strategy.type" : "DefaultRolloverStrategy",
```

```
    "appender.main.strategy.max" : "10"  
  },  
}  
]
```

## Amazon EMR バージョン 5.29.0 以降

Amazon EMR バージョン 5.29.0 以前の場合、Flink では log4j ロギングフレームワークが使用されません。次の設定例は、log4j 形式を示しています。

```
[  
  {  
    "Classification": "flink-log4j",  
    "Properties": {  
      "log4j.appender.file": "org.apache.log4j.RollingFileAppender",  
      "log4j.appender.file.append": "true",  
      # keep up to 4 files and each file size is limited to 100MB  
      "log4j.appender.file.MaxFileSize": "100MB",  
      "log4j.appender.file.MaxBackupIndex": 4,  
      "log4j.appender.file.layout": "org.apache.log4j.PatternLayout",  
      "log4j.appender.file.layout.ConversionPattern": "%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p  
%-60c %x - %m%n"  
    },  
  }  
]
```

## Flink が Java 11 で実行されるよう設定する

Amazon EMR リリース 6.12.0 以降では、Java 11 ランタイム環境で Flink を実行できます。このセクションでは、Java 11 ランタイム環境で Flink を実行するための設定について説明します。

### トピック

- [Java 11 で実行する Flink をクラスター作成時に設定する](#)
- [稼働しているクラスターの Java 11 で実行されている Flink を設定する](#)
- [稼働中のクラスターで Flink の Java ランタイムを確認します。](#)

### Java 11 で実行する Flink をクラスター作成時に設定する

Flink を Java 11 ランタイムで実行する EMR クラスターを作成するには、次の手順に従います。Java 11 ランタイム環境を使用できるようにする設定ファイルは `flink-conf.yaml` です。

## New console

Flink を Java 11 ランタイム環境で実行するクラスターを新しいコンソールで作成するには

1. にサインインし AWS Management Console、<https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. ナビゲーションペインの [EMR on EC2] で [クラスター] を選択し、[クラスターの作成] を選択します。
3. Amazon EMR リリース 6.12.0 以降を選択し、Flink アプリケーションのインストールを選択します。クラスターにインストールする他のアプリケーションを選択します。
4. クラスターの設定をそのまま進めます。オプションの [ソフトウェア設定] セクションで、デフォルトの [設定の入力] オプションを使用して、次の設定を入力します。

```
[
  {
    "Classification": "flink-conf",
    "Properties": {
      "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
      "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
      "env.java.home":"/usr/lib/jvm/jre-11"
    }
  }
]
```

5. 続けてクラスターの設定と起動を行います。

## AWS CLI

Flink を Java 11 ランタイム環境で実行するクラスターを CLI から作成するには

1. Flink を Java 11 で実行する設定ファイル `configurations.json` を作成する

```
[
  {
    "Classification": "flink-conf",
    "Properties": {
      "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
      "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
      "env.java.home":"/usr/lib/jvm/jre-11"
    }
  }
]
```

```
}  
]
```

2. から AWS CLI、Amazon EMR リリース 6.12.0 以降で新しい EMR クラスターを作成し、次の例に示すように Flink アプリケーションをインストールします。

```
aws emr create-cluster --release-label emr-6.12.0 \  
--applications Name=Flink \  
--configurations file://./configurations.json \  
--region us-east-1 \  
--log-uri s3://myLogUri \  
--instance-type m5.xlarge \  
--instance-count 2 \  
--service-role EMR_DefaultRole_V2 \  
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

## 稼働しているクラスターの Java 11 で実行されている Flink を設定する

Flink を Java 11 ランタイムで実行している稼働中の EMR クラスターを更新するには、次の手順に従います。Java 11 ランタイム環境を使用できるようにする設定ファイルは `flink-conf.yaml` です。

### New console

Flink を Java 11 ランタイムで実行している稼働中のクラスターを新しいコンソールで更新するには

1. にサインインし AWS Management Console、<https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. ナビゲーションペインの [EMR on EC2] で [クラスター] を選択し、更新するクラスターを選択します。

#### Note

クラスターで Java 11 をサポートするには、Amazon EMR リリース 6.12.0 以降を使用する必要があります。

3. [設定] タブを選択します。

4. [インスタンスグループ設定] セクションで、更新する [実行中] のインスタンスグループを選択し、リストアクションメニューから [再設定] を選択します。
5. [属性の編集] オプションを使用してインスタンスグループを次のように再設定します。各設定の後に、[新しい設定の追加] を選択します。

分類	プロパティ	値
flink-conf	containerized.taskmanager.env.JAVA_HOME	/usr/lib/jvm/jre-11
flink-conf	containerized.master.env.JAVA_HOME	/usr/lib/jvm/jre-11
flink-conf	env.java.home	/usr/lib/jvm/jre-11

6. [変更の保存] を選択して設定を追加します。

## AWS CLI

Flink を Java 11 ランタイムで実行している稼働中のクラスターを CLI から更新するには

`modify-instance-groups` コマンドを使用して、実行中のクラスター内のインスタンスグループに新しい設定を指定します。

1. 最初に、設定ファイル `configurations.json` を作成し、Java 11 を使用できるようにします。次の例にある `ig-1xxxxxxx9` を、再設定するインスタンスグループの ID に置き換えます。そのファイルを、`modify-instance-groups` コマンドを実行するディレクトリに保存します。

```
[
  {
    "InstanceGroupId": "ig-1xxxxxxx9",
    "Configurations": [
      {
        "Classification": "flink-conf",
        "Properties": {
          "containerized.taskmanager.env.JAVA_HOME": "/usr/lib/jvm/jre-11",
```



```
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
    },
    "Configurations":[]
}
]
}
```

2. から AWS CLI、次のコマンドを実行します。再設定するインスタンスグループの ID を置き換えます。

```
aws emr modify-instance-groups --cluster-id j-2AL4XXXXXX5T9 \
--instance-groups file://configurations.json
```

稼働中のクラスターで Flink の Java ランタイムを確認します。

稼働中のクラスターの Java ランタイムを確認するには、「[SSH を使用してプライマリノードに接続する](#)」の説明どおり、プライマリノードに SSH でログインします。次に、以下のコマンドを実行します。

```
ps -ef | grep flink
```

ps コマンドで `-ef` オプションを指定すると、システム上で実行中のプロセスをすべて一覧表示できます。その出力を `grep` でフィルタリングすると、`flink` の文字列が含まれる行を検索できます。Java ランタイム環境 (JRE) の値 `jre-XX` を出力で確認します。次の出力の `jre-11` は、Flink のランタイム環境として Java 11 が選択されることを示しています。

```
flink    19130    1  0 09:17 ?        00:00:15 /usr/lib/jvm/jre-11/bin/
java -Djava.io.tmpdir=/mnt/tmp -Dlog.file=/usr/lib/flink/log/flink-flink-
historyserver-0-ip-172-31-32-127.log -Dlog4j.configuration=file:/usr/lib/flink/conf/
log4j.properties -Dlog4j.configurationFile=file:/usr/lib/flink/conf/log4j.properties
-Dlogback.configurationFile=file:/usr/lib/flink/conf/logback.xml -classpath /usr/lib/
flink/lib/flink-cep-1.17.0.jar:/usr/lib/flink/lib/flink-connector-files-1.17.0.jar:/
usr/lib/flink/lib/flink-csv-1.17.0.jar:/usr/lib/flink/lib/flink-json-1.17.0.jar:/usr/
lib/flink/lib/flink-scala_2.12-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-java-
uber-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-scala-bridge_2.12-1.17.0.
```

または、[プライマリノードに SSH でログイン](#)し、コマンド `flink-yarn-session -d` を使用して Flink YARN セッションを開始します。次の例の出力には、Flink の Java Virtual Machine (JVM) として `java-11-amazon-corretto` が表示されています。

```
2023-05-29 10:38:14,129 INFO org.apache.flink.configuration.GlobalConfiguration
  [] - Loading configuration property: containerized.master.env.JAVA_HOME, /usr/lib/
  jvm/java-11-amazon-corretto.x86_64
```

## Amazon EMR 内で Flink ジョブを操作する

Amazon EMR で Flink を操作するには、コンソール、ResourceManager 追跡 UI にある Flink インターフェイス、コマンドラインを使用するという複数の方法があります。このいずれの場合にも、JAR ファイルを Flink アプリケーションに送信できます。JAR ファイルを送信すると、Flink によって管理されるジョブになります JobManager。JobManager は、Flink セッションアプリケーションマスターデーモンをホストする YARN ノードにあります。

長期実行のクラスターまたは一時的なクラスターで、YARN ジョブとして Flink アプリケーションを実行できます。長期実行のジョブでは、複数の Flink ジョブを Amazon EMR で実行する 1 つの Flink クラスターに送信できます。一時的なクラスターで Flink ジョブを稼働する場合、Amazon EMR クラスターは Flink アプリケーションを実行するために必要な時間のためだけに存在します。そのため、使用したリソースと費やした時間に対してのみ課金されます。Flink ジョブを送信するには、Amazon EMR AddSteps API オペレーションによって RunJobFlow オペレーションのステップを調整するか、AWS CLI の `add-steps` または `create-cluster` コマンドを使用します。

## 長期実行のクラスターのステップとして、Flink YARN アプリケーションを起動します

複数のクライアントが YARN API オペレーションを介して作業を送信できる Flink アプリケーションを起動するには、クラスターを作成するか、既存のクラスターに Flink アプリケーションを追加する必要があります。新しいクラスターを作成する手順については、「[Flink を使用してクラスターを作成する](#)」を参照してください。既存のクラスターで YARN セッションを開始するには、コンソール、AWS CLI、Java SDK から次のステップに従います。

### Note

Amazon EMR バージョン 5.5.0 では、実行を簡素化するための `yarn-session.sh` スクリプトのラッパーとして、`flink-yarn-session` コマンドが追加されました。以前のバージョンの Amazon EMR を使用している場合は、コンソールの引数、または AWS CLI コマン

この Args を `bash -c "/usr/lib/flink/bin/yarn-session.sh -d"` で置き換えます。

コンソールを使用して Flink ジョブを既存のクラスターに送信するには

`flink-yarn-session` コマンドを使用して、Flink セッションを既存のクラスターに送信します。

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. クラスターリストで、以前に起動したクラスターを選択します。
3. クラスターの詳細ページで、[Steps (ステップ)]、[Add Step (ステップの追加)] の順に選択します。
4. 次のガイドラインに従ってパラメータを入力し、[追加] を選択します。

パラメータ	説明
ステップタイプ	カスタム JAR
名前	ステップを識別するのに役立つ名前。例えば、 <code>#example-flink-step-name#</code> などです。
Jar location	command-runner.jar
引数	<code>flink-yarn-session</code> コマンドとアプリケーションに適切な引数。例えば、 <code>flink-yarn-session -d</code> は YARN クラスター内でデタッチされた状態 () で Flink セッションを開始します-d。引数の詳細については、最新の Flink ドキュメントの「 <a href="#">YARN Setup (YARN の設定)</a> 」を参照してください。

AWS CLI を使用して Flink ジョブを既存のクラスターに送信するには

- Flink ジョブを長期実行クラスターに追加するには、`add-steps` コマンドを使用します。次のコマンド例では、YARN クラスター内において Flink セッションが切り離された状態 (-d) で開

始されるよう、`Args="flink-yarn-session", "-d"` を指定しています。引数の詳細については、最新の Flink ドキュメントの「[YARN Setup \(YARN の設定\)](#)」を参照してください。

```
aws emr add-steps --cluster-id <j-XXXXXXXX> --steps Type=CUSTOM_JAR,Name=<example-fLink-step-name>,Jar=command-runner.jar,Args="flink-yarn-session","-d"
```

## 長期実行クラスター上の既存の Flink アプリケーションに作業を送信する

長期実行クラスターに既存の Flink アプリケーションがある場合は、作業を送信するために、クラスターの Flink アプリケーション ID を指定できます。アプリケーション ID を取得するには、`yarn application -list` AWS CLI または [YarnClient](#) API オペレーションを使用して を実行します。

```
$ yarn application -list
16/09/07 19:32:13 INFO client.RMProxy: Connecting to ResourceManager at
ip-10-181-83-19.ec2.internal/10.181.83.19:8032
Total number of applications (application-types: [] and states: [SUBMITTED, ACCEPTED,
RUNNING]):1
Application-Id      Application-Name      Application-Type      User      Queue      State
Final-State      Progress      Tracking-URL
application_1473169569237_0002      Flink session with 14 TaskManagers (detached)
Apache Flink      hadoop      default      RUNNING      UNDEFINED
100% http://ip-10-136-154-194.ec2.internal:33089
```

この Flink セッションのアプリケーション ID は `application_1473169569237_0002` であり、これを使用することで、AWS CLI または SDK からアプリケーションに操作内容を送信できます。

### Example SDK for Java

```
List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("flink", "run", "-m", "yarn-cluster", "-yid",
"application_1473169569237_0002", "-yn", "2", "/usr/lib/flink/examples/streaming/
WordCount.jar",
    "--input", "s3://myBucket/pg11.txt", "--output", "s3://myBucket/alice2/");

StepConfig flinkRunWordCount = new StepConfig()
    .withName("Flink add a wordcount step")
    .withActionOnFailure("CONTINUE")
    .withHadoopJarStep(flinkWordCountConf);
```

```
stepConfigs.add(flinkRunWordCount);

AddJobFlowStepsResult res = emr.addJobFlowSteps(new AddJobFlowStepsRequest()
    .withJobFlowId("myClusterId")
    .withSteps(stepConfigs));
```

## Example AWS CLI

```
aws emr add-steps --cluster-id <j-XXXXXXXX> \
--steps Type=CUSTOM_JAR,Name=Flink_Submit_To_Long_Running,Jar=command-runner.jar,\
Args="flink","run","-m","yarn-cluster","-yid","application_1473169569237_0002",\
"/usr/lib/flink/examples/streaming/WordCount.jar",\
"--input","s3://myBucket/pg11.txt","--output","s3://myBucket/alice2/" \
--region <region-code>
```

## 一時的な Flink ジョブを送信する

次の例では Flink ジョブを実行する一時的なクラスターを起動し、完了時に終了します。

## Example SDK for Java

```
import java.util.ArrayList;
import java.util.List;
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;

public class Main_test {

    public static void main(String[] args) {
        AWSCredentials credentials_profile = null;
        try {
            credentials_profile = new ProfileCredentialsProvider("default").getCredentials();
        } catch (Exception e) {
            throw new AmazonClientException(
                "Cannot load credentials from .aws/credentials file. " +
                "Make sure that the credentials file exists and the profile name is specified within it.",
            );
        }
    }
}
```

```
e);
}

AmazonElasticMapReduce emr = AmazonElasticMapReduceClientBuilder.standard()
    .withCredentials(new AWSStaticCredentialsProvider(credentials_profile))
    .withRegion(Regions.US_WEST_1)
    .build();

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();
HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("bash", "-c", "flink", "run", "-m", "yarn-cluster", "-yn", "2",
        "/usr/lib/flink/examples/streaming/WordCount.jar", "--input", "s3://path/to/
input-file.txt", "--output",
        "s3://path/to/output/");

StepConfig flinkRunWordCountStep = new StepConfig()
    .withName("Flink add a wordcount step and terminate")
    .withActionOnFailure("CONTINUE")
    .withHadoopJarStep(flinkWordCountConf);

stepConfigs.add(flinkRunWordCountStep);

Application flink = new Application().withName("Flink");

RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("flink-transient")
    .withReleaseLabel("emr-5.20.0")
    .withApplications(flink)
    .withServiceRole("EMR_DefaultRole")
    .withJobFlowRole("EMR_EC2_DefaultRole")
    .withLogUri("s3://path/to/my/logfiles")
    .withInstances(new JobFlowInstancesConfig()
        .withEc2KeyName("myEc2Key")
        .withEc2SubnetId("subnet-12ab3c45")
        .withInstanceCount(3)
        .withKeepJobFlowAliveWhenNoSteps(false)
        .withMasterInstanceType("m4.large")
        .withSlaveInstanceType("m4.large"))
    .withSteps(stepConfigs);

RunJobFlowResult result = emr.runJobFlow(request);
System.out.println("The cluster ID is " + result.toString());
```

```
}  
  
}
```

## Example AWS CLI

Flink ジョブの完了時に終了する一時的なクラスターを作成するには、`create-cluster` サブコマンドを使用します。

```
aws emr create-cluster --release-label emr-5.2.1 \  
--name "Flink_Transient" \  
--applications Name=Flink \  
--configurations file:///./configurations.json \  
--region us-east-1 \  
--log-uri s3://myLogUri \  
--auto-terminate \  
--instance-type m5.xlarge \  
--instance-count 2 \  
--service-role EMR_DefaultRole_V2 \  
--ec2-attributes KeyName=<YourKeyName>,InstanceProfile=EMR_EC2_DefaultRole \  
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\ \  
Args="bash","-c","\\"flink run -m yarn-cluster /usr/lib/flink/examples/streaming/\ \  
WordCount.jar \  
--input s3://myBucket/pg11.txt --output s3://myBucket/alice/""
```

## Scala シェルを使用する

EMR クラスターの Flink Scala シェルは、YARN の新規セッションを開始する目的でのみ設定できます。次の手順に従って Scala シェルを使用できます。

プライマリノードで Flink Scala シェルを使用する

1. 「[SSH を使用してプライマリノードに接続する](#)」の説明どおり、プライマリノードに SSH でログインします。
2. 次のように入力してシェルを開始します。

Amazon EMR バージョン 5.5.0 以降では、次のコマンドを使用して、1 つの TaskManager で Scala シェルの Yarn クラスターを起動できます。

```
% flink-scala-shell yarn 1
```

以前のバージョンの Amazon EMR では、以下を使用します。

```
% /usr/lib/flink/bin/start-scala-shell.sh yarn 1
```

これで Flink Scala を起動して、インタラクティブに Flink を使用できます。他のインターフェイスやオプションと同様に、シェルから起動するタスクの数に応じて、例に使用されている `-n` オプションの値を増やせます。

詳細については、Apache Flink の公式ドキュメントの「[Scala REPL](#)」を参照してください。

## Flink ウェブインターフェイスを検索する

Flink ウェブインターフェイスは、Flink アプリケーションホストに属するアプリケーションマスターによってホストされています。このインターフェイスによっても、ジョブとしての JAR を送信したり、他のジョブについて現在のステータスを表示したりできます。Flink ウェブインターフェイスは、実行中の Flink セッションがあるかぎりにはアクティブです。長期実行 YARN ジョブが既にアクティブになっている場合は、「Amazon EMR 管理ガイド」の「[SSH を使用してプライマリノードに接続する](#)」トピックの手順に従って、YARN ResourceManager に接続できます。たとえば、SSH トンネルを設定していて、ブラウザでプロキシを有効にしている場合、EMR クラスターの詳細ページで、[Connections] (接続) に ResourceManager 接続を選択します。

Cluster: Development Cluster Waiting Cluster ready after last step completed.

Connections:



Resource Manager ... (View All)

ResourceManager を見つけたら、Flink セッションをホストしている YARN アプリケーションを選択します。[Tracking UI (UI 追跡)] 列でリンクを選択します。



# All Applications

Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealed Nodes
2	2 GB	11.25 GB	0 B	2	8	0	1	0	0	0	0

Scheduling Resource Type	Minimum Allocation	Maximum Allocation
MEMORY	<memory:32, vCores:1>	<memory:11520, vCores:8>

Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI
Flink session with 1 TaskManagers (detached)	Apache Flink	default	Mon Oct 10 14:42:47 -0700 2016	N/A	RUNNING	UNDEFINED		<a href="#">ApplicationMaster</a>

First Previous

Flink ウェブインターフェイスでは、設定を表示して、独自のカスタム JAR をジョブとして送信したり、実行中のジョブをモニタリングしたりできます。

The screenshot shows the Apache Flink Dashboard interface. On the left is a navigation menu with options: Overview, Running Jobs, Completed Jobs, Task Managers, Job Manager, and Submit new Job. The main content area is titled 'Overview' and shows 'Version: 1.1.1'. It features three summary cards: '1 Task Managers', '1 Task Slots', and '1 Available Task Slots'. To the right is a 'Total Jobs' summary table:

Total Jobs	Count
Running	0
Finished	0
Canceled	0
Failed	0

Below these are two tables for 'Running Jobs' and 'Completed Jobs', each with columns for Start Time, End Time, Duration, Job Name, Job ID, Tasks, and Status.

## Flink オートスケーラー

### 概要

Amazon EMR リリース 6.15.0 以降では、Flink オートスケーラーがサポートされています。ジョブ オートスケーラー機能は、実行中の Flink ストリーミングジョブからメトリクスを収集し、個々の

ジョブ頂点を自動的にスケールリングします。これにより、バックプレッシャーが軽減され、設定した使用率目標を達成できます。

詳細については、「[Apache Flink Kubernetes Operator ドキュメント](#)」の「[オートスケーラー](#)」セクションを参照してください。

## 考慮事項

- Flink オートスケーラーは、Amazon EMR 6.15.0 以降でサポートされています。
- Flink Autoscaler は、ストリーミングジョブでのみサポートされています。
- アダプティブスケジューラーのみがサポートされます。デフォルトのスケジューラーはサポートされていません。
- 動的なリソースプロビジョニングを可能にするために、クラスタースケールリングを有効化することを推奨します。メトリクスの評価は 5~10 秒ごとに行われるため、Amazon EMR Managed Scaling が推奨されます。この間隔では、クラスターは必要なクラスターリソースの変化により簡単に適応できます。

## オートスケーラーを有効にする

EC2 クラスターで Amazon EMR を作成するときに、次のステップを使用して Flink オートスケーラーを有効にします。

1. Amazon EMR コンソールで、EMR クラスターを新規作成します。
  - a. Amazon EMR リリース `emr-6.15.0` 以降を選択します。[Flink] アプリケーションバンドルを選択し、クラスターに含めるその他のアプリケーションを選択します。

### Name and applications [Info](#)








**Name**

**Amazon EMR release** [Info](#)

A release contains a set of applications which can be installed on your cluster.

emr-6.15.0 ▼

**Application bundle**

<p>Spark Interactive</p> 	<p>Core Hadoop</p> 	<p>Flink</p> 	<p>HBase</p> 	<p>Presto</p> 	<p>Trino</p> 	<p>Custom</p> 
--	--	--	--	--	--	---

<input checked="" type="checkbox"/> Flink 1.17.1	<input type="checkbox"/> Ganglia 3.7.2	<input type="checkbox"/> HBase 2.4.17
<input type="checkbox"/> HCatalog 3.1.3	<input checked="" type="checkbox"/> Hadoop 3.3.6	<input checked="" type="checkbox"/> Hive 3.1.3
<input type="checkbox"/> Hue 4.11.0	<input type="checkbox"/> JupyterEnterpriseGateway 2.6.0	<input type="checkbox"/> JupyterHub 1.5.0
<input type="checkbox"/> Livy 0.7.1	<input type="checkbox"/> MXNet 1.9.1	<input type="checkbox"/> Oozie 5.2.1
<input type="checkbox"/> Phoenix 5.1.3	<input type="checkbox"/> Pig 0.17.0	<input type="checkbox"/> Presto 0.283
<input type="checkbox"/> Spark 3.4.1	<input type="checkbox"/> Sqoop 1.4.7	<input type="checkbox"/> TensorFlow 2.11.0
<input type="checkbox"/> Tez 0.10.2	<input type="checkbox"/> Trino 426	<input checked="" type="checkbox"/> Zeppelin 0.10.1
<input type="checkbox"/> ZooKeeper 3.5.10		

- b. [クラスターのスケールリングとプロビジョニング] で [EMR マネージドスケールリングを使用] を選択します。

### Cluster scaling and provisioning [Info](#)

Set up scaling and provisioning configurations for the core and task node groups for your cluster.

Choose an option

<input type="radio"/> <b>Set cluster size manually</b> Use this option if you know your workload patterns in advance.	<input checked="" type="radio"/> <b>Use EMR-managed scaling</b> Monitor key workload metrics so that EMR can optimize the cluster size and resource utilization.	<input type="radio"/> <b>Use custom automatic scaling</b> To programmatically scale core and task nodes, create custom automatic scaling policies.
--	---	---

2. [ソフトウェア設定] セクションで、次の設定を入力して Flink オートスケーラーを有効にします。テストシナリオでは、検証を容易にするために、決定間隔、メトリクスウィンドウ間隔、および安定化間隔を低い値に設定し、ジョブがすぐにスケールリングを決定できるようにします。

```
[
  {
    "Classification": "flink-conf",
```

```
"Properties": {
  "job.autoscaler.enabled": "true",
  "jobmanager.scheduler": "adaptive",
  "job.autoscaler.stabilization.interval": "60s",
  "job.autoscaler.metrics.window": "60s",
  "job.autoscaler.decision.interval": "10s",
  "job.autoscaler.debug.logs.interval": "60s"
}
}
```

- 必要に応じて他の設定を選択または構成し、Flink オートスケーラー対応クラスターを作成します。

## オートスケーラーの設定

このセクションでは、特定のニーズに基づいて変更できる構成のうちほとんどについて説明します。

### Note

time、interval、window 設定のような時間ベースの構成では、単位が指定されていない場合のデフォルト単位はミリ秒です。そのため、サフィックスなしの 30 の値は 30 ミリ秒になります。他の時間単位には、秒には s、分には m、時間には h という適切なサフィックスを含めてください。

### トピック

- [オートスケーラーのループ設定](#)
- [メトリクスの集計と履歴の設定](#)
- [ジョブ頂点レベルの設定](#)
- [バックログ処理設定](#)
- [スケール操作設定](#)

## オートスケーラーのループ設定

オートスケーラーは、設定可能な数回の時間間隔ごとにジョブ頂点レベルのメトリクスを取得してスケールアクションに変換し、新しいジョブ頂点の並列処理を推定して、ジョブスケジューラーに推奨

します。メトリクスは、ジョブの再起動時間とクラスターの安定化間隔が過ぎた後にのみ収集されません。

設定キー	デフォルト値	説明	値の例
<code>job.autoscaler.enabled</code>	<code>false</code>	Flink クラスターでオートスケーリングを有効にします。	<code>true</code> , <code>false</code>
<code>job.autoscaler.decision.interval</code>	<code>60s</code>	オートスケーラーの決定間隔。	30 (デフォルト単位はミリ秒)、5m、1h
<code>job.autoscaler.restart.time</code>	<code>3m</code>	オペレーターが履歴から確実に再起動を判断できるようになるまでの予想される再起動時間。	30 (デフォルト単位はミリ秒)、5m、1h
<code>job.autoscaler.stabilization.interval</code>	<code>300s</code>	新しいスケーリングが実行されない安定化期間。	30 (デフォルト単位はミリ秒)、5m、1h
<code>job.autoscaler.debug.logs.interval</code>	<code>300s</code>	オートスケーラーのデバッグログの間隔。	30 (デフォルト単位はミリ秒)、5m、1h

## メトリクスの集計と履歴の設定

オートスケーラーはメトリクスを取得し、時間ベースのスライディングウィンドウに沿って集計し、評価してスケーリングを決定します。各ジョブ頂点のスケーリング決定履歴は、新しい並列処理の見積もりに利用されます。これらには、時間ベースの有効期限と履歴サイズ (少なくとも 1) の両方があります。

設定キー	デフォルト値	説明	値の例
<code>job.autoscaler.metrics.window</code>	<code>600s</code>	Scaling metrics aggregation window size.	30 (デフォルト単位はミリ秒)、5m、1h

設定キー	デフォルト値	説明	値の例
job.autoscaler.history.max.count	3	頂点ごとに保持できる過去のスケール決定の最大数。	1 ~ Integer.MAX_VALUE
job.autoscaler.history.max.age	24h	頂点ごとに保持する過去のスケール決定の最小数。	30 (デフォルト単位はミリ秒)、5m、1h

## ジョブ頂点レベルの設定

各ジョブの頂点の並列処理は、ターゲットの使用率に基づいて変更され、最小/最大並列処理の制限によって制限されます。目標使用率を 100% に近い値 (つまり 1) に設定することは推奨されず、使用率境界は中間の負荷変動を処理するバッファの役割を果たします。

設定キー	デフォルト値	説明	値の例
job.autoscaler.target.utilization	0.7	目標とする頂点使用率。	0 - 1
job.autoscaler.target.utilization.boundary	0.4	目標とする頂点利用率境界。現在の処理速度が $[\text{target\_rate} / (\text{target\_utilization} - \text{boundary})]$ および $(\text{target\_rate} / (\text{target\_utilization} + \text{boundary}))$ 以内の場合、スケールは実行されません。	0 - 1
job.autoscaler.vertex.min-parallelism	1	オートスケーラーが使用できる最低限の並列処理。	0 - 200
job.autoscaler.vertex.max-parallelism	200	オートスケーラーが使用できる最大並列処理。この制限は、Flink 構成または各オペレ	0 - 200

設定キー	デフォルト値	説明	値の例
		ーターで直接構成された最大並列処理よりも高い場合、無視されることに注意してください。	

## バックログ処理設定

ジョブ頂点には、スケール操作期間中に蓄積される保留中のイベント、つまりバックログを処理するための追加リソースが必要です。これは catch-up 期間とも呼ばれます。バックログの処理時間が設定された lag -threshold 値を超えると、ジョブ頂点ターゲットの使用率は最大レベルまで増加します。これにより、バックログの処理中に不要なスケールアップ操作が行われるのを防ぐことができます。

設定キー	デフォルト値	説明	値の例
job.autoscaler.backlog-processing.lag-threshold	5m	ラグのしきい値。ラグの原因となる保留中のメッセージを削除しながら、不要なスケールアップを防止します。	30 (デフォルト単位はミリ秒)、5m、1h
job.autoscaler.catch-up.duration	15m	スケールアップ操作後にバックログを完全に処理するまでの目標時間。0 に設定すると、バックログベースのスケールアップが無効になります。	30 (デフォルト単位はミリ秒)、5m、1h

## スケール操作設定

オートスケーラーは、猶予期間内のスケールアップ操作の直後には、スケールダウン操作を実行しません。これにより、一時的な負荷変動によって発生する、スケールアップ、スケールダウンが繰り返し発生する不要なサイクルを防ぐことができます。

スケールダウン操作比率を利用して並列処理を徐々に減らし、一時的な負荷の急上昇に対応するためにリソースを解放することができます。また、大規模なスケールダウンの後に不要なマイナーなスケールアップ操作が行われるのを防ぐのにも役立ちます。

過去のジョブ頂点スケーリング決定履歴に基づいて無効なスケールアップ操作を検出し、さらなる並列処理の変更を防ぐことができます。

設定キー	デフォルト値	説明	値の例
<code>job.autoscaler.scale-up.grace-period</code>	1h	頂点をスケールアップした後に、その頂点をスケールダウンできない期間。	30 (デフォルト単位はミリ秒)、5m、1h
<code>job.autoscaler.scale-down.max-factor</code>	0.6	最大スケールダウン係数。1の値は、スケールダウンに制限がなく、0.6 はジョブを元の並列処理の 60% でのみスケールダウンできることを意味します。	0 - 1
<code>job.autoscaler.scale-up.max-factor</code>	100000.	最大スケールアップ率。2.0の値は、現在の並列処理の 200% でのみジョブをスケールアップできることを意味します。	0 - Integer.MAX_VALUE
<code>job.autoscaler.scaling.effectiveness.detection.enabled</code>	false	効果のないスケーリング操作の検出を有効にし、オートスケーラーが今後のスケールアップをブロックできるようにするかどうか。	true, false

## タスク回復とスケーリング操作のためのジョブ再起動時間の最適化

タスクが失敗したり、スケーリング操作が発生したりすると、Flink は最後に完了したチェックポイントからタスクを再実行しようとしています。チェックポイントの状態のサイズと並列タスクの数によっては、再起動プロセスの実行に 1 分以上かかる場合があります。再起動中は、ジョブのバックログタスクが蓄積されることがあります。ただし、Flink が実行グラフの回復と再開の速度を最適化してジョブの安定性を高める方法はいくつかあります。



このページでは、Amazon EMR Flink がタスクリカバリまたはスケーリング操作中のジョブの再起動時間を改善できるいくつかの方法について説明します。

## トピック

- [タスクローカルリカバリ](#)
- [汎用ログベースのインクリメンタルチェックポイント](#)
- [きめ細かなリカバリ](#)
- [アダプティブスケジューラーに組み込まれた再起動メカニズム](#)

## タスクローカルリカバリ

### Note

タスクローカルリカバリは、Amazon EMR 6.0.0 以降でサポートされています。

Flink チェックポイントでは、各タスクが Flink が Amazon S3 などの分散ストレージに書き込む状態のスナップショットを作成します。復旧時には、タスクは分散ストレージから状態を復元します。分散ストレージはすべてのノードからアクセスできるため、耐障害性があり、再スケーリング中に状態を再分散できます。

ただし、リモート分散ストアには欠点もあります。すべてのタスクはネットワーク経由でリモートロケーションから状態を読み取る必要があります。そのため、タスクの回復やスケーリング操作中に大きな状態の回復時間が長くなる可能性があります。

リカバリ時間が長いというこの問題は、タスクローカルリカバリによって解決できます。タスクはチェックポイントの状態を、ローカルディスクなど、タスクのローカルにあるセカンダリストレージに書き込みます。また、プライマリストレージ（この場合は Amazon S3）に状態が保存されます。復元中、スケジューラーは、タスクが以前に実行されていたのと同じタスクマネージャー上でタスクをスケジュールし、リモート状態ストアから読み取るのではなく、ローカル状態ストアから復元できるようにします。詳細については、「[Apache Flink ドキュメント](#)」の「[タスクローカルリカバリ](#)」を参照してください。

サンプルジョブを使ったベンチマークテストでは、タスクローカルリカバ리를有効にすると、リカバリ時間が数分から数秒に短縮されたことがわかりました。

タスクローカルリカバ리를有効にするには、`flink-conf.yaml` ファイルで次の設定を行います。チェックポイント間隔の値をミリ秒単位で指定します。

```
state.backend.local-recovery: true
state.backend: hasmap or rocksdb
state.checkpoints.dir: s3://storage-location-bucket-path/checkpoint
execution.checkpointing.interval: 15000
```

## 汎用ログベースのインクリメンタルチェックポイント

### Note

Amazon EMR 6.10.0 以降では、汎用ログベースのインクリメンタルチェックポイントがサポートされています。

チェックポイントの速度を向上させるため、汎用ログベースのインクリメンタルチェックポイントが Flink 1.16 に追加されました。チェックポイント間隔を短くすると、回復後に再処理する必要のあるイベントが少なくなるため、多くの場合回復作業が削減されます。詳細については、「[Apache Flink ブログ](#)」の「[汎用ログベースのインクリメンタルチェックポイントによるチェックポイントの速度と安定性の向上](#)」を参照してください。

サンプルジョブを使用したベンチマークテストでは、汎用ログベースのインクリメンタルチェックポイントを使用すると、チェックポイントにかかる時間が数分から数秒に短縮されたことがわかりました。

汎用ログベースのインクリメンタルチェックポイントを有効にするには、`flink-conf.yaml` ファイルで次の設定を行います。チェックポイント間隔の値をミリ秒単位で指定します。

```
state.backend.changelog.enabled: true
state.backend.changelog.storage: filesystem
dst1.dfs.base-path: s3://bucket-path/changelog
state.backend.local-recovery: true
state.backend: rocksdb
state.checkpoints.dir: s3://bucket-path/checkpoint
execution.checkpointing.interval: 15000
```

## きめ細かなリカバリ

### Note

Amazon EMR 6.0.0 以降では、デフォルトスケジューラーのきめ細かな復元がサポートされています。Amazon EMR 6.15.0 以降では、アダプティブスケジューラーでのきめ細かなリカバリのサポートを利用できます。

実行中にタスクが失敗した場合、Flink は実行グラフ全体をリセットし、最後に完了したチェックポイントから完全な再実行をトリガーします。これは、失敗したタスクを単に再実行するよりもコストがかかります。きめ細かい復元では、失敗したタスクのパイプラインに接続されたコンポーネントのみを再起動します。次の例では、ジョブグラフには 5 つの頂点 (A から E) があります。頂点間のすべての接続はポイントごとの分散でパイプライン化され、ジョブの `parallelism.default` は 2 に設定されます。

```
A # B # C # D # E
```

この例では、合計 10 個のタスクが実行されています。最初のパイプライン (a1 から e1) は TaskManager (TM1) で実行され、2 番目のパイプライン (a2 から e2) は別の TaskManager (TM2) 上で実行されます。

```
a1 # b1 # c1 # d1 # e1  
a2 # b2 # c2 # d2 # e2
```

パイプライン接続されたコンポーネントには、a1 # e1 と a2 # e2 の 2 つがあります。TM1 または TM2 のどちらか一方で障害が発生しても、影響を受けるのは、その TaskManager が実行されていたパイプライン内の 5 つのタスクのみです。再起動戦略では、影響を受けるパイプラインコンポーネントのみを起動します。

きめ細かいリカバリは、完全に並列した Flink ジョブでのみ機能します。keyBy() または redistribute() オペレーションではサポートされていません。詳細については、「Flink 改善提案」Jira プロジェクトの「[FLIP-1: タスク障害からのきめ細かな回復](#)」を参照してください。

きめ細かい復元を有効にするには、flink-conf.yaml ファイルで次の設定を行います。

```
jobmanager.execution.failover-strategy: region  
restart-strategy: exponential-delay or fixed-delay
```

## アダプティブスケジューラーに組み込まれた再起動メカニズム

### Note

アダプティブスケジューラーの複合再起動メカニズムは、Amazon EMR 6.15.0 以降でサポートされています。

アダプティブスケジューラーは、使用可能なスロットに基づいてジョブの並列処理を調整できます。設定したジョブの並列処理を満たすだけの十分なスロットがない場合は、自動的に並列処理を減らします。新しいスロットが使用可能になると、ジョブは設定されたジョブの並列処理に合わせて再びスケールアップされます。適応型スケジューラーは、利用可能なリソースが十分になくともジョブのダウンタイムを回避します。これは Flink Autoscaler でサポートされているスケジューラーです。これらの理由から、Amazon EMR Flink を使用するアダプティブスケジューラーをお勧めします。ただし、アダプティブスケジューラーは、新しいリソースが追加されるたびに 1 回再起動するなど、短期間に複数の再起動を行う場合があります。これにより、ジョブのパフォーマンスが低下する可能性があります。

Amazon EMR 6.15.0 以降では、Flink のアダプティブスケジューラーに再起動メカニズムが組み合わされています。このメカニズムは、最初のリソースが追加されると再起動ウィンドウを開き、設定したウィンドウ間隔（デフォルトの 1 分）まで待機します。並列処理を設定してジョブを実行するのに十分なリソースがあるとき、または間隔がタイムアウトになったときに、1 回再起動します。

サンプルジョブを使用したベンチマークテストでは、アダプティブスケジューラーと Flink オートスケジューラーを使用すると、この機能はデフォルトの動作よりも 10% 多くのレコードを処理することがわかりました。

複合再起動メカニズムを有効にするには、`flink-conf.yaml` ファイルで以下の設定を行います。

```
jobmanager.adaptive-scheduler.combined-restart.enabled: true
jobmanager.adaptive-scheduler.combined-restart.window-interval: 1m
```

## Amazon EMR の Zeppelin から Flink ジョブを操作する

### 序章

Amazon EMR リリース 6.10.0 以降では、[Apache Zeppelin](#) と Apache Flink の連携がサポートされており、Flink ジョブを Zeppelin ノートブックからインタラクティブに送信できます。Flink イ

インタープリタを使用すると、Flink クエリの実行、Flink ストリーミングおよびバッチのジョブ定義、Zeppelin ノートブック内での出力の視覚化を行えます。Flink インタープリタは、Flink REST API をベースに構築されています。これにより、Zeppelin 環境内から Flink ジョブにアクセスして操作し、データの処理と分析をリアルタイムに実行できます。

Flink インタープリタには 4 つのサブインタープリターがあります。これらは、異なる目的を持ちますが、すべて JVM 内に存在し、事前設定された、Flink への同じエントリポイント (ExecutionEnvironment、StreamExecutionEnvironment、BatchTableEnvironment、StreamTableEnvironment) を共有しています。具体的なインタープリタは、次のとおりです。

- `%flink` — ExecutionEnvironment、StreamExecutionEnvironment、BatchTableEnvironment、StreamTableEnvironment を作成し、Scala 環境を提供する
- `%flink.pyflink` — Python 環境を提供する
- `%flink.ssql` — ストリーミング SQL 環境を提供する
- `%flink.bsql` — バッチ SQL 環境を提供する

## 前提条件

- Zeppelin と Flink の連携は、Amazon EMR 6.10.0 以降で作成されたクラスターでサポートされています。
- これらの手順に応じて EMR クラスターでホストされているウェブインターフェイスを表示するには、SSH トンネルを設定してインバウンドアクセスを許可する必要があります。詳細については、「[プライマリノードでホストされるウェブサイトを表示するようにプロキシを設定する](#)」を参照してください。

## EMR クラスターで Zeppelin と Flink の連携を設定する

Apache Zeppelin で動作する Apache Flink が EMR クラスター上で稼働できるよう設定するには、次の手順に従います。


1. Amazon EMR コンソールからクラスターを新規作成します。emr-6.10.0 以降の Amazon EMR リリースを選択します。次に、[カスタム] オプションを使用して、アプリケーションバンドルのカスタマイズを選択します。バンドルには、少なくとも Flink、Hadoop、Zeppelin を含めてください。


Amazon EMR release [Info](#)


A release contains a set of applications which can be installed on your cluster.


emr-6.10.0 ▼


Application bundle


Spark  


Core  
Hadoop  


HBase  


Presto  


Trino  


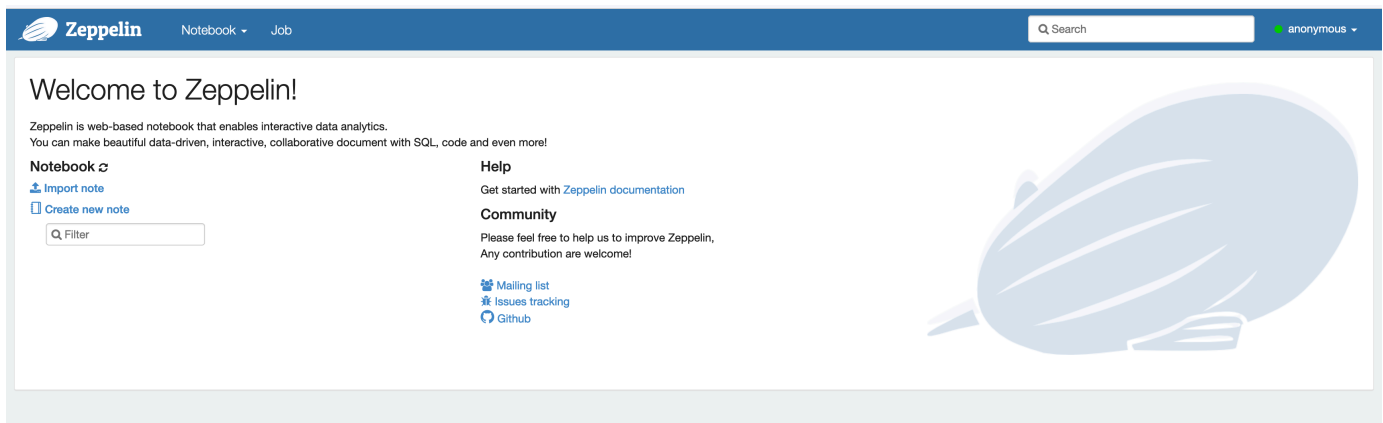
Custom  


▼ Customize your application bundle

Applications included in bundle

<input checked="" type="checkbox"/> Flink 1.16.0	<input type="checkbox"/> Ganglia 3.7.2
<input type="checkbox"/> HBase 2.4.15	<input type="checkbox"/> HCatalog 3.1.3
<input checked="" type="checkbox"/> Hadoop 3.3.3	<input type="checkbox"/> Hive 3.1.3
<input type="checkbox"/> Hue 4.10.0	<input type="checkbox"/> JupyterEnterpriseGateway 2.6.0
<input type="checkbox"/> JupyterHub 1.5.0	<input type="checkbox"/> Livy 0.7.1
<input type="checkbox"/> MXNet 1.9.1	<input type="checkbox"/> Oozie 5.2.1
<input type="checkbox"/> Phoenix 5.1.2	<input type="checkbox"/> Pig 0.17.0
<input type="checkbox"/> Presto 0.278	<input type="checkbox"/> Spark 3.3.1
<input type="checkbox"/> Sqoop 1.4.7	<input type="checkbox"/> TensorFlow 2.11.0
<input type="checkbox"/> Tez 0.10.2	<input type="checkbox"/> Trino 403
<input checked="" type="checkbox"/> Zeppelin 0.10.1	<input type="checkbox"/> ZooKeeper 3.5.10

- 残りのクラスターは、任意の設定で作成します。
- クラスターを稼働させたら、そのクラスターをコンソールから選択して詳細を表示し、[アプリケーション] タブを開きます。[アプリケーションユーザーインターフェイス] セクションから [Zeppelin] を選択して、Zeppelin ウェブインターフェイスを開きます。[前提条件](#)の説明どおり、プライマリノードへの SSH トンネルとプロキシ接続を介して、Zeppelin ウェブインターフェイスにアクセスできるように設定されていることを確認します。



Zeppelin

Notebook - Job

Q Search anonymous ▼

Welcome to Zeppelin!

Zeppelin is web-based notebook that enables interactive data analytics.  
You can make beautiful data-driven, interactive, collaborative document with SQL, code and even more!

**Notebook** ▾

↳ Import note

□ Create new note

Q Filter

**Help**

Get started with [Zeppelin documentation](#)

**Community**

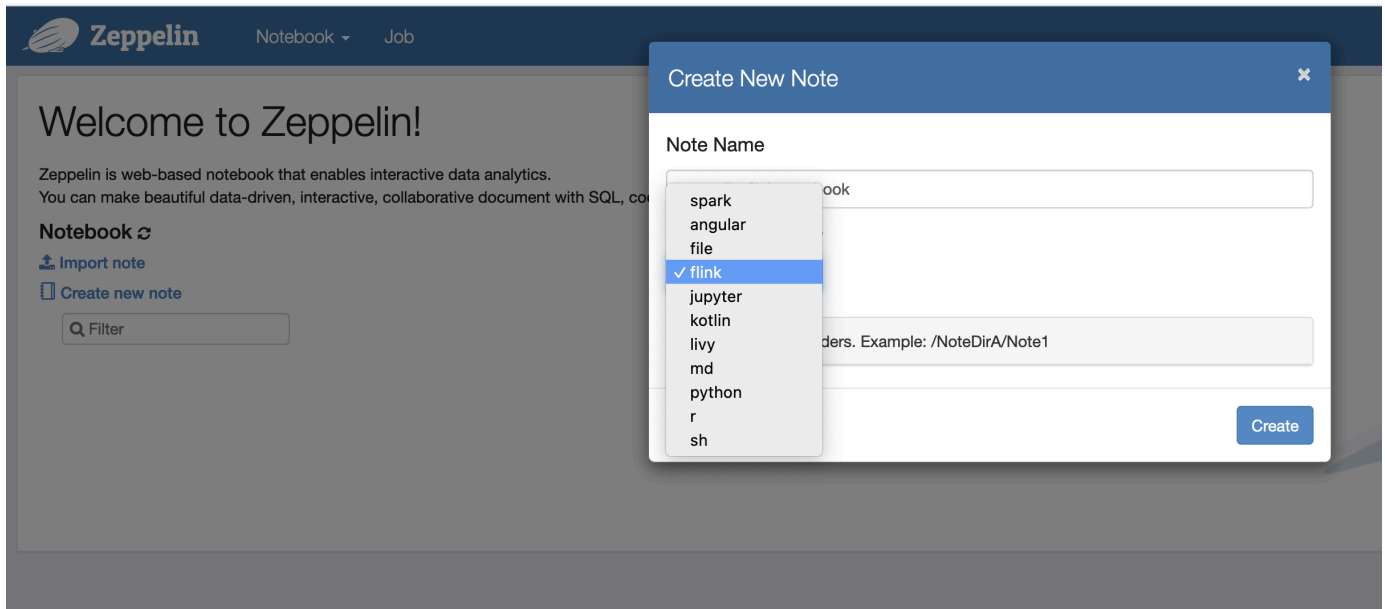
Please feel free to help us to improve Zeppelin,  
Any contribution are welcome!

✉ Mailing list

🔗 Issues tracking

🔗 Github

4. これで、Flink をデフォルトのインタープリタとして使用して、Zeppelin ノートブックにノート を新規作成できるようになりました。



5. Zeppelin ノートブックから Flink ジョブを実行する方法については、次のコード例を参照してください。

## Zeppelin と Flink が連携する EMR クラスターで Flink ジョブを実行する

### • 例 1: Flink Scala

#### a) バッチ WordCount の例 (SCALA)

```
%flink

val data = benv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
    .map(w => (w, 1))
    .groupBy(0)
    .sum(1)
    .print()
```

#### b) ストリーミング WordCount の例 (SCALA)

```
%flink

val data = senv.fromElements("hello world", "hello flink", "hello hadoop")
```

```
data.flatMap(line => line.split("\\s"))
  .map(w => (w, 1))
  .keyBy(0)
  .sum(1)
  .print

senv.execute()
```

**Batch WordCount** FLINK JOB FINISHED

```
%flink
val data = benv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
  .map(w => (w, 1))
  .groupBy(0)
  .sum(1)
  .print()

data: org.apache.flink.api.scala.DataSet[String] = org.apache.flink.api.scala.DataSet@22fe7dd5
(flink,1)
(hadoop,1)
(hello,3)
(world,1)

Took 56 sec. Last updated by anonymous at May 04 2023, 2:19:24 PM. (updated)
```

**Streaming WordCount** FLINK JOB FINISHED

```
%flink
val data = senv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
  .map(w => (w, 1))
  .keyBy(0)
  .sum(1)
  .print

senv.execute()

date: org.apache.flink.streaming.api.scala.DataStream[String] = org.apache.flink.streaming.api.scala.DataStream@282340f2
warning: there was one deprecation warning; for details, enable ':setting -deprecation' or ':replay -deprecation'
res2: org.apache.flink.streaming.api.datastream.DataStreamSink[(String, Int)] = org.apache.flink.streaming.api.d
atostream.DataStreamSink@7a0e7b2f
res3: org.apache.flink.api.common.JobExecutionResult =
Program execution finished
Job with JobID a78d1a9b33c0fb71b8d53f00cda51030 has finished.
Job Runtime: 8485 ms

Took 12 sec. Last updated by anonymous at May 04 2023, 2:21:05 PM. (updated)
```

## • 例 2: Flink ストリーミング SQL

```
%flink.sql
SET 'sql-client.execution.result-mode' = 'tableau';
SET 'table.dml-sync' = 'true';
SET 'execution.runtime-mode' = 'streaming';

create table dummy_table (
  id int,
  data string
) with (
  'connector' = 'filesystem',
  'path' = 's3://s3-bucket/dummy_table',
  'format' = 'csv'
);

INSERT INTO dummy_table SELECT * FROM (VALUES (1, 'Hello World'), (2, 'Hi'), (2,
'Hi'), (3, 'Hello'), (3, 'World'), (4, 'ADD'), (5, 'LINE'));

SELECT * FROM dummy_table;
```



```
%flink.sql
SET 'sql-client.execution.result-mode' = 'tableau';
SET 'table.dml-sync' = 'true';
SET 'execution.runtime-mode' = 'streaming';

create table t1 (
  id int,
  data string
) with (
  'connector' = 'filesystem',
  'path' = 's3://dbsamrat-emr-dev/glue-catalog-test/dbsamrat/t1/',
  'format' = 'csv'
);

INSERT INTO t1 SELECT * FROM (VALUES (1, 'Hello World'), (2, 'Hi'), (2, 'Hi'), (3, 'Hello'), (3, 'World'), (4, 'ADD'), (5, 'LINE'));
SELECT * FROM t1;
```

id	data
3	Hello
3	World
3	Hello
3	World
4	ADD
4	ADD
5	LINE
5	LINE

Took 20 sec. Last updated by anonymous at May 04 2023, 3:07:03 PM. (outdated)

- 例 3: Pyflink。という名前の独自のサンプルテキストファイルを S3 word.txt バケットにアップロードする必要があることに注意してください。

```
%flink.pyflink
```

```
import argparse
import logging
import sys
```

```
from pyflink.common import Row
from pyflink.table import (EnvironmentSettings, TableEnvironment, TableDescriptor,
                           Schema,
                           DataTypes, FormatDescriptor)
from pyflink.table.expressions import lit, col
from pyflink.table.udf import udtf
```

```
def word_count(input_path, output_path):
    t_env = TableEnvironment.create(EnvironmentSettings.in_streaming_mode())
    # write all the data to one file
    t_env.get_config().set("parallelism.default", "1")

    # define the source
    if input_path is not None:
        t_env.create_temporary_table(
            'source',
            TableDescriptor.for_connector('filesystem')
```

```

        .schema(Schema.new_builder()
                .column('word', DataTypes.STRING())
                .build())
        .option('path', input_path)
        .format('csv')
        .build()
    tab = t_env.from_path('source')
else:
    print("Executing word_count example with default input data set.")
    print("Use --input to specify file input.")
    tab = t_env.from_elements(map(lambda i: (i,), word_count_data),
                              DataTypes.ROW([DataTypes.FIELD('line',
DataTypes.STRING())]))

# define the sink
if output_path is not None:
    t_env.create_temporary_table(
        'sink',
        TableDescriptor.for_connector('filesystem')
            .schema(Schema.new_builder()
                    .column('word', DataTypes.STRING())
                    .column('count', DataTypes.BIGINT())
                    .build())
            .option('path', output_path)
            .format(FormatDescriptor.for_format('canal-json')
                    .build())
            .build())
else:
    print("Printing result to stdout. Use --output to specify output path.")
    t_env.create_temporary_table(
        'sink',
        TableDescriptor.for_connector('print')
            .schema(Schema.new_builder()
                    .column('word', DataTypes.STRING())
                    .column('count', DataTypes.BIGINT())
                    .build())
            .build())

@udtf(result_types=[DataTypes.STRING()])
def split(line: Row):
    for s in line[0].split():
        yield Row(s)

# compute word count

```

```
tab.flat_map(split).alias('word') \  
  .group_by(col('word')) \  
  .select(col('word'), lit(1).count) \  
  .execute_insert('sink') \  
  .wait()
```

```
logging.basicConfig(stream=sys.stdout, level=logging.INFO, format="%(message)s")
```

```
word_count("s3://s3_bucket/word.txt", "s3://s3_bucket/demo_output.txt")
```

1. Zeppelin UI で [FLINK JOB] を選択すると、Flink ウェブの UI にアクセスし、表示を行えます。



**Batch WordCount**

FLINK JOB FINISHED

```
%flink  
  
val data = benv.fromElements("hello world", "hello flink", "hello hadoop")  
data.flatMap(line => line.split("\\s"))  
  .map(w => (w, 1))  
  .groupBy(0)  
  .sum(1)  
  .print()  
  
data: org.apache.flink.api.scala.DataSet[String] = org.apache.flink.api.scala.DataSet@22fe7dd5  
(flink,1)  
(hadoop,1)  
(hello,3)  
(world,1)  
  
Took 56 sec. Last updated by anonymous at May 04 2023, 2:19:24 PM. (outdated)
```

2. [FLINK JOB] を選択すると、Flink ウェブコンソールにルーティングされ、そのページがブラウザの別のタブで開きます。

The screenshot shows the Apache Flink Dashboard interface. On the left is a navigation sidebar with options: Overview, Jobs (selected), Running Jobs, Completed Jobs, Task Managers, Job Manager, and Submit New Job. The main content area displays:

- Available Task Slots:** 0. Total Task Slots: 0, Task Managers: 0.
- Running Jobs:** 0. Finished: 2, Canceled: 0, Failed: 0.
- Running Job List:** A table with columns: Job Name, Start Time, Duration, End Time, Tasks, Status. It shows "No Data".
- Completed Job List:** A table with columns: Job Name, Start Time, Duration, End Time, Tasks, Status.
 

Job Name	Start Time	Duration	End Time	Tasks	Status
Flink Streaming Job	2023-05-04 14:20:56	8s	2023-05-04 14:21:04	2 / 2	FINISHED
Flink Java Job at Thu May 04 08:49:10 UTC 2023	2023-05-04 14:19:11	12s	2023-05-04 14:19:23	3 / 3	FINISHED

## Flink リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Flink のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Flink バージョン情報

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.36.2	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
		rn-timeline-server, flink-client, flink-jobmanager-config
emr-7.1.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.0.0	1.18.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-6.15.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.14.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-6.13.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.12.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-6.11.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.11.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.10.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi



Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-6.10.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.9.1	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.9.0	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-6.8.1	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.8.0	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.7.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.36.1	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.36.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.6.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.35.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.5.0	1.14.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.4.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-6.3.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.3.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.2.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-6.2.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.1.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-6.1.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.34.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.33.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.33.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.32.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.32.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.31.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config



Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.31.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.30.2	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.30.1	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.30.0	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.29.0	1.9.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.28.1	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.28.0	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.27.1	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.27.0	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.26.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.25.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.24.1	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.24.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.23.1	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.23.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.22.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya rn-timeline-server, flink-client
emr-5.21.2	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya rn-timeline-server, flink-client
emr-5.21.1	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya rn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.21.0	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.20.1	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.20.0	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.19.1	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.19.0	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.18.1	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client



Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.18.0	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.2	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.1	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.17.0	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.16.1	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.16.0	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.15.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.15.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.2	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.14.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.13.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.13.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.3	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.2	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.12.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.4	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.11.3	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.2	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.11.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.10.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.10.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client



Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.9.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.9.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.3	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.8.2	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.1	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.0	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.7.1	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.7.0	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.6.1	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.6.0	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya rn-timeline-server, flink-client
emr-5.5.4	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.3	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.5.2	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.4.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.4.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.2	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.3.1	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.0	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.3	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.2.2	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client



Amazon EMR リリースラベル	Flink バージョン	Flink でインストールされるコンポーネント
emr-5.1.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.1.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

# Ganglia

## Note

Ganglia を含む Amazon EMR の最後のリリースは Amazon EMR 6.15.0 でした。クラスターをモニタリングするために、6.15.0 以降のリリースには [Amazon CloudWatch エージェント](#) が含まれています。

Ganglia オープンソースプロジェクトは拡張可能な分散システムで、パフォーマンスへの影響を最小限に抑えながら、クラスターやグリッドをモニタリングできるように設計されています。クラスターで Ganglia を有効にすると、レポートを生成し、クラスター全体のパフォーマンスを表示するだけでなく、個別のノードインスタンスのパフォーマンスを調べることができます。また、Ganglia は、Hadoop および Spark メトリクスを取り込み、視覚化するように設定されています。Ganglia オープンソースプロジェクトの詳細については、<http://ganglia.info/> を参照してください。

ブラウザで Ganglia ウェブ UI を表示すると、クラスターのパフォーマンスの概要が表示されます。これらのレポートには、クラスターのロード、メモリの使用状況、CPU 使用率、ネットワークトラフィックなどがグラフで詳しく示されています。クラスターの統計情報の下には、クラスター内の個別のサーバーごとにグラフがあります。

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Ganglia のバージョンと、Amazon EMR で Ganglia と共にインストールされるコンポーネントを示しています。

このリリースで Ganglia と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の Ganglia バージョン情報

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.15.0	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
		server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, webserver

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Ganglia のバージョンと、Amazon EMR で Ganglia と共にインストールされるコンポーネントを示しています。

このリリースで Ganglia と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

#### emr-5.36.2 の Ganglia バージョン情報

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.36.2	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, webserver

#### トピック

- [Ganglia を使用したクラスターの作成](#)
- [Ganglia メトリクスを表示する](#)
- [Ganglia での Hadoop と Spark のメトリクス](#)
- [Ganglia リリース履歴](#)

# Ganglia を使用したクラスターの作成

## Note

Ganglia を含む Amazon EMR の最後のリリースは Amazon EMR 6.15.0 でした。クラスターをモニタリングするために、6.15.0 以降のリリースには [Amazon CloudWatch エージェント](#) が含まれています。

コンソールを使用し、Ganglia でクラスターを作成するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [クラスターを作成] を選択します。
3. [Software configuration (ソフトウェア設定)] で、[All Applications (すべてのアプリケーション)]、[Core Hadoop]、または [Spark] を選択します。
4. 適切な設定でクラスターの作成を進めます。

を使用してクラスターに Ganglia を追加するには AWS CLI

では AWS CLI、`--applications` パラメータ `create-cluster` でを使用して Ganglia をクラスターに追加できます。`--applications` パラメータを使用して Ganglia のみを指定する場合、Ganglia が、インストールされる唯一のアプリケーションです。

- クラスターを作成するときに、次のコマンドを入力して Ganglia を追加し、*myKey* を EC2 キーペアの名前に置き換えます。

## Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Spark cluster with Ganglia" --release-label  
emr-6.15.0; \
```

```
--applications Name=Spark Name=Ganglia \  
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \  
--instance-count 3 --use-default-roles
```

--instance-groups パラメータを使用せずにインスタンス数を指定すると、1つのマスターノードが起動され、残りのインスタンスはコアノードとして起動されます。すべてのノードで、コマンドで指定したインスタンスタイプが使用されます。

#### Note

以前にデフォルトの EMR サービスロールと EC2 インスタンスプロファイルを作成していない場合は、「aws emr create-default-roles」と入力してそれらを作成してから、create-cluster サブコマンドを入力します。

での Amazon EMR コマンドの使用の詳細については AWS CLI、「」を参照してください <https://docs.aws.amazon.com/cli/latest/reference/emr>。

## Ganglia メトリクスを表示する

#### Note

Ganglia を含む Amazon EMR の最後のリリースは Amazon EMR 6.15.0 でした。クラスターをモニタリングするために、6.15.0 以降のリリースには [Amazon CloudWatch エージェント](#) が含まれています。

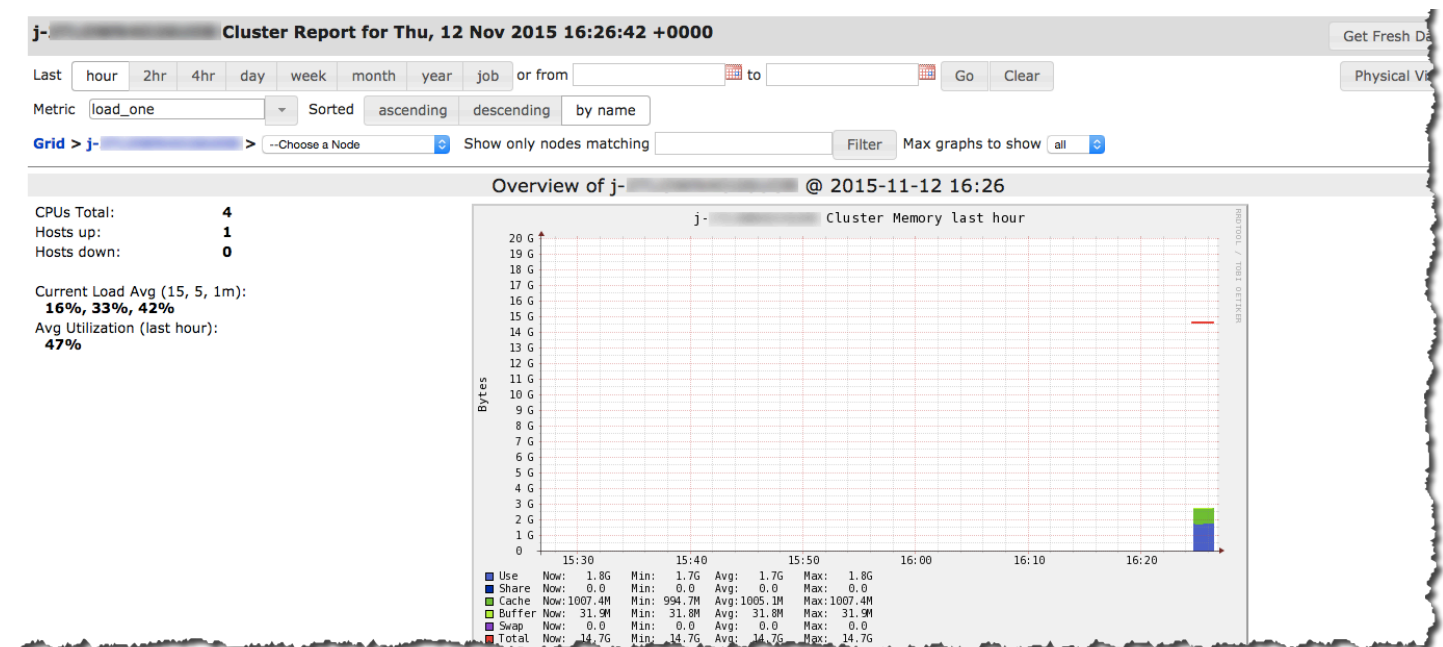
Ganglia には、ウェブベースのユーザーインターフェイスが用意されており、Ganglia が収集するメトリクスは、このインターフェイスを使用して表示できます。Amazon EMR で Ganglia を実行するとき、ウェブインターフェイスはマスターノードで実行され、ポート転送を使用して表示できます。これは、SSH トンネルの作成とも呼ばれます。Amazon EMR のウェブインターフェイスの表示の詳細については、「Amazon EMR 管理ガイド」の「[EMR クラスターでホストされているウェブインターフェイスの表示](#)」を参照してください。

Ganglia のウェブインターフェイスを表示するには

1. SSH を使用してマスターノードにトンネル接続し、安全な接続を確立します。マスターノードへの SSH トンネルの作成方法の詳細については、「Amazon EMR 管理ガイド」の「[オプショ](#)

[ン 2、パート 1: 動的ポートフォワーディングを使用してマスターノードへの SSH トンネルをセットアップする](#)」を参照してください。

- Firefox 用 FoxyProxy プラグインなどのプロキシツールを使用してウェブブラウザをインストールし、\*ec2\*.amazonaws.com\* タイプのドメインの SOCKS プロキシを作成します。詳細については、「[Amazon EMR 管理ガイド](#)」の「オプション 2、パート 2: マスターノードでホストされるウェブサイトを表示するようにプロキシを設定する」を参照してください。
- プロキシを設定し、SSH 接続を開いた状態で、`http://master-public-dns-name/ganglia/` でブラウザウィンドウを開くことで Ganglia UI を表示できます。ここで、`master-public-dns-name` は EMR クラスター内のマスターサーバーのパブリック DNS アドレスです。



## Ganglia での Hadoop と Spark のメトリクス

### Note

Ganglia を含む Amazon EMR の最後のリリースは Amazon EMR 6.15.0 でした。クラスターをモニタリングするために、6.15.0 以降のリリースには [Amazon CloudWatch エージェント](#) が含まれています。

Ganglia は、各インスタンスに対して Hadoop メトリクスをレポートします。さまざまなタイプのメトリクスには、分散ファイルシステム (dfs.\*)、Java 仮想マシン (jvm.\*)、MapReduce (mapred.\*)、リモートプロシージャコール (rpc.\*) のカテゴリごとにプレフィックスが付けられます。

Spark や Hadoop などの YARN ベースの Ganglia メトリクスは、EMR リリースバージョン 4.4.0 および 4.5.0 では使用できません。これらのメトリクスを使用するには、それ以降のバージョンを使用してください。

通常、Spark の Ganglia メトリクスには YARN アプリケーション ID および Spark DAGScheduler 用のプレフィックスがあります。したがって、プレフィックスは次の形式になります。

- DAGScheduler.\*
- application\_XXXXXXXXXX\_XXXX.driver.\*
- application\_XXXXXXXXXX\_XXXX.executor.\*

## Ganglia リリース履歴

### Note

Ganglia を含む Amazon EMR の最後のリリースは Amazon EMR 6.15.0 でした。クラスターをモニタリングするために、6.15.0 以降のリリースには [Amazon CloudWatch エージェント](#) が含まれています。

次の表は、Amazon EMR の各リリースに含まれている Ganglia のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Ganglia バージョン情報

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.36.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
		client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver



Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.36.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.36.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.35.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, webserver
emr-6.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, webserver
emr-6.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, webserver



Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-6.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.34.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.33.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.33.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.31.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.31.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.30.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.29.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.28.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.28.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.27.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.27.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.26.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.25.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.24.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.24.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.23.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.23.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.22.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver



Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.21.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.20.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.20.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.18.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.18.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.17.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.16.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.16.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.14.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.13.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.12.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver



Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.11.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.5.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.5.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-5.5.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-5.5.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver



Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.3.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-5.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-5.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.2.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-5.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-5.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.0.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-5.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-4.9.6	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-4.9.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-4.9.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-4.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-4.8.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-4.8.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-4.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver



Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-4.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-4.7.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver
emr-4.7.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-4.7.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-4.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR リリースラベル	Ganglia バージョン	Ganglia でインストールされるコンポーネント
emr-4.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.2.0	3.6.0	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

# Apache Hadoop

[Apache Hadoop はオープンソースの Java ソフトウェアフレームワークで、インスタンスクラスター全体で大量のデータ処理をサポートします。](#) インスタンス 1 つから数千個まで規模を問わずに動作します。Hadoop は、MapReduce や Tez などのさまざまな処理モデルを使用して複数のインスタンスに処理を分散し、HDFS と呼ばれる分散ファイルシステムを使用して複数のインスタンスにデータを保存します。Hadoop はクラスター内の各インスタンスの稼働状況を監視し、1 つまたは複数のノードが障害を起こしても自動的に復帰させます。このように、Hadoop は処理能力やストレージ性能を向上させるだけでなく、高い可用性も提供するものです。詳細については、[Hadoop ドキュメント](#)を参照してください。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Hadoop のバージョンと、Amazon EMR で Hadoop と共にインストールされるコンポーネントを示しています。

このリリースで Hadoop と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Hadoop バージョン情報

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-7.1.0	Hadoop 3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Hadoop のバージョンと、Amazon EMR で Hadoop と共にインストールされるコンポーネントを示しています。

このリリースで Hadoop と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Hadoop バージョン情報

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.15.0	Hadoop 3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Hadoop のバージョンと、Amazon EMR で Hadoop と共にインストールされるコンポーネントを示しています。

このリリースで Hadoop と共にインストールされるコンポーネントのバージョンについては、「[リリース 5.36.2 コンポーネントバージョン](#)」を参照してください。

#### emr-5.36.2 の Hadoop バージョン情報

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.36.2	Hadoop 2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
		nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 5.18.0 以降では、Amazon EMR アーティファクトリポジトリを使用し、特定の Amazon EMR リリースで利用可能なライブラリおよび依存関係の正確なバージョンに対してジョブコードをビルドできます。詳細については、「[Amazon EMR アーティファクトリポジトリを使用して依存関係を確認する](#)」を参照してください。

## トピック

- [Hadoop の設定](#)
- [Amazon EMR における HDFS での透過的暗号化](#)
- [Hadoop アプリケーションの作成または実行](#)
- [YARN コンテナの不均一なメモリアクセス対応を有効にする](#)
- [Hadoop のバージョン履歴](#)

## Hadoop の設定

以下のセクションでは、Hadoop デーモン、タスク、および HDFS 用のデフォルトの設定を行います。

### トピック

- [タスクの設定](#)
- [Hadoop デーモン構成設定](#)
- [HDFS 構成](#)

## タスクの設定

設定変数を設定して、ジョブのパフォーマンスを調整できます MapReduce。このセクションでは、重要な設定のデフォルト値を提供します。デフォルト値は、クラスターで使用されるノードの EC2 インスタンスタイプにより異なります。HBase は Amazon EMR リリースバージョン 4.6.0 以降を使

用するときには使用できます。HBase がインストールされているときは、別のデフォルトが使用されます。これらの値は、初期のデフォルトと共に示します。

Hadoop 2 は 2 つのパラメータ、`mapreduce.map.java.opts` および `mapreduce.reduce.java.opts` を使用して、マップとリデュース JVM のメモリをそれぞれ設定します。これらのパラメータは、Hadoop バージョンの 1 つの設定オプション (`mapreduce.map.java.opts`) に置き換わります。

同様に、Hadoop 2.7.2 以降では `mapred.job.jvm.num.tasks` が `mapred.job.reuse.jvm.num.tasks` に置き換わります。Amazon EMR では、EC2 インスタンスタイプにかかわらず、この値を 20 に設定します。`mapred-site` 設定分類を使用して、この設定を上書きできます。-1 の値を設定すると、JVM を単一のジョブ内の無限のタスクで再利用でき、1 の値は、新しい JVM が各タスクにスポンされることを示します。

たとえば、`mapred.job.jvm.num.tasks` の値を -1 に設定すると、次の内容でファイルを作成することができます。

```
[
  {
    "Classification": "mapred-site",
    "Properties": {
      "mapred.job.jvm.num.tasks": "-1"
    }
  }
]
```

から `create-cluster` コマンドまたは `modify-instance-groups` コマンドを使用すると AWS CLI、JSON 設定ファイルを参照できます。次の例では、設定ファイルは `myConfig.json` として保存され、Amazon S3 に格納されます。

#### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --release-label emr-7.1.0 --instance-type m5.xlarge \  
--instance-count 3 --applications Name=Hadoop --configurations https://  
s3.amazonaws.com/mybucket/myfolder/myConfig.json \  

```



```
--use-default-roles
```

同様に、`mapred-site` 設定分類を使用して、以下のデフォルト値を変更することができ、単一の JSON ファイルを使用して、複数の値と複数の設定分類を設定できます。詳細については、「[アプリケーションの設定](#)」を参照してください。

Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## タスクの構成設定のデフォルト値

### インスタンスタイプ

- [c1 インスタンス](#)
- [c3 インスタンス](#)
- [c4 インスタンス](#)
- [c5 インスタンス](#)
- [c5a インスタンス](#)
- [c5ad インスタンス](#)
- [c5d インスタンス](#)
- [c5n インスタンス](#)
- [c6a インスタンス](#)
- [c6g インスタンス](#)
- [c6gd インスタンス](#)
- [c6gn インスタンス](#)
- [c6i インスタンス](#)
- [c6id インスタンス](#)
- [c6in インスタンス](#)
- [c7a インスタンス](#)
- [c7g インスタンス](#)
- [c7gd インスタンス](#)

- [c7gn インスタンス](#)
- [c7i インスタンス](#)
- [d2 インスタンス](#)
- [d3 インスタンス](#)
- [d3en インスタンス](#)
- [g3 インスタンス](#)
- [g3s インスタンス](#)
- [g4dn インスタンス](#)
- [g5 インスタンス](#)
- [h1 インスタンス](#)
- [i2 インスタンス](#)
- [i3 インスタンス](#)
- [i3en インスタンス](#)
- [i4g インスタンス](#)
- [i4i インスタンス](#)
- [im4gn インスタンス](#)
- [is4gen インスタンス](#)
- [m1 インスタンス](#)
- [m2 インスタンス](#)
- [m3 インスタンス](#)
- [m4 インスタンス](#)
- [m5 インスタンス](#)
- [m5a インスタンス](#)
- [m5ad インスタンス](#)
- [m5d インスタンス](#)
- [m5dn インスタンス](#)
- [m5n インスタンス](#)
- [m5zn インスタンス](#)
- [m6a インスタンス](#)
- [m6g インスタンス](#)

- [m6gd インスタンス](#)
- [m6i インスタンス](#)
- [m6id インスタンス](#)
- [m6idn インスタンス](#)
- [m6in インスタンス](#)
- [m7a インスタンス](#)
- [m7g インスタンス](#)
- [m7gd インスタンス](#)
- [m7i インスタンス](#)
- [m7i-flex インスタンス](#)
- [p2 インスタンス](#)
- [p3 インスタンス](#)
- [p5 インスタンス](#)
- [r3 インスタンス](#)
- [r4 インスタンス](#)
- [r5 インスタンス](#)
- [r5a インスタンス](#)
- [r5ad インスタンス](#)
- [r5b インスタンス](#)
- [r5d インスタンス](#)
- [r5dn インスタンス](#)
- [r5n インスタンス](#)
- [r6a インスタンス](#)
- [r6g インスタンス](#)
- [r6gd インスタンス](#)
- [r6i インスタンス](#)
- [r6id インスタンス](#)
- [r6idn インスタンス](#)
- [r6in インスタンス](#)
- [r7a インスタンス](#)

- [r7g インスタンス](#)
- [r7gd インスタンス](#)
- [r7i インスタンス](#)
- [r7iz インスタンス](#)
- [x1 インスタンス](#)
- [x1e インスタンス](#)
- [x2gd インスタンス](#)
- [x2idn インスタンス](#)
- [x2iedn インスタンス](#)
- [z1d インスタンス](#)

## c1 インスタンス

### c1.medium

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx288m</code>	<code>-Xmx288m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx288m</code>	<code>-Xmx288m</code>
<code>mapreduce.map.memory.mb</code>	512	512
<code>mapreduce.reduce.memory.mb</code>	512	512
<code>yarn.app.mapreduce.am.resource.mb</code>	512	512
<code>yarn.scheduler.minimum-allocation-mb</code>	256	256
<code>yarn.scheduler.maximum-allocation-mb</code>	512	512
<code>yarn.nodemanager.resource.memory-mb</code>	1024	512

## c1.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1024	1024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	256	256
yarn.scheduler.maximum-allocation-mb	2048	2560
yarn.nodemanager.resource.memory-mb	5120	2560

## c3 インスタンス

## c3.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c3.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

## c3.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

## c3.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1331m	-Xmx1331m
mapreduce.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.map.memory.mb	1664	1664
mapreduce.reduce.memory.mb	3328	3328
yarn.app.mapreduce.am.resource.mb	3328	3328

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	53248	26624
<code>yarn.nodemanager.resource.memory-mb</code>	53248	26624

## c4 インスタンス

### c4.large

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx717m</code>	<code>-Xmx717m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx1434m</code>	<code>-Xmx1434m</code>
<code>mapreduce.map.memory.mb</code>	896	896
<code>mapreduce.reduce.memory.mb</code>	1792	1792
<code>yarn.app.mapreduce.am.resource.mb</code>	1792	1792
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	1792	896
<code>yarn.nodemanager.resource.memory-mb</code>	1792	896



## c4.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c4.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11520	5760
<code>yarn.nodemanager.resource.memory-mb</code>	11520	5760

## c4.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1152m</code>	<code>-Xmx1152m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2304m</code>	<code>-Xmx2304m</code>
<code>mapreduce.map.memory.mb</code>	1440	1440
<code>mapreduce.reduce.memory.mb</code>	2880	2880
<code>yarn.app.mapreduce.am.resource.mb</code>	2880	2880
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23040	11520
<code>yarn.nodemanager.resource.memory-mb</code>	23040	11520

## c4.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1183m	-Xmx1183m
mapreduce.java.opts	-Xmx2366m	-Xmx2366m
mapreduce.map.memory.mb	1479	1479
mapreduce.reduce.memory.mb	2958	2958
yarn.app.mapreduce.am.resource.mb	2958	2958
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

## c5 インスタンス

## c5.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

## c5.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

## c5.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## c5.9xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	65536	32768
<code>yarn.nodemanager.resource.memory-mb</code>	65536	32768

## c5.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1502m</code>	<code>-Xmx1502m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3004m</code>	<code>-Xmx3004m</code>
<code>mapreduce.map.memory.mb</code>	1877	1877
<code>mapreduce.reduce.memory.mb</code>	3754	3754
<code>yarn.app.mapreduce.am.resource.mb</code>	3754	3754
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	90112	30048
<code>yarn.nodemanager.resource.memory-mb</code>	90112	30048

## c5.18xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1547m	-Xmx1547m
mapreduce.java.opts	-Xmx3094m	-Xmx3094m
mapreduce.map.memory.mb	1934	1934
mapreduce.reduce.memory.mb	3868	3868
yarn.app.mapreduce.am.resource.mb	3868	3868
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

## c5.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	188416	31376
<code>yarn.nodemanager.resource.memory-mb</code>	188416	31376

## c5a インスタンス

### c5a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1126m</code>	<code>-Xmx1126m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2252m</code>	<code>-Xmx2252m</code>
<code>mapreduce.map.memory.mb</code>	1408	1408
<code>mapreduce.reduce.memory.mb</code>	2816	2816
<code>yarn.app.mapreduce.am.resource.mb</code>	2816	2816
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	5632	2816
<code>yarn.nodemanager.resource.memory-mb</code>	5632	2816



## c5a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c5a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23424	11712
<code>yarn.nodemanager.resource.memory-mb</code>	23424	11712

## c5a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1357m</code>	<code>-Xmx1357m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2714m</code>	<code>-Xmx2714m</code>
<code>mapreduce.map.memory.mb</code>	1696	1696
<code>mapreduce.reduce.memory.mb</code>	3392	3392
<code>yarn.app.mapreduce.am.resource.mb</code>	3392	3392
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c5a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

## c5a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	116736	29184
<code>yarn.nodemanager.resource.memory-mb</code>	116736	29184

## c5a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1494m</code>	<code>-Xmx1494m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2988m</code>	<code>-Xmx2988m</code>
<code>mapreduce.map.memory.mb</code>	1867	1867
<code>mapreduce.reduce.memory.mb</code>	3734	3734
<code>yarn.app.mapreduce.am.resource.mb</code>	3734	3734
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29840
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29840

## c5ad インスタンス

## c5ad.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c5ad.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c5ad.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c5ad.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## c5ad.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074

## c5ad.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1459m</code>	<code>-Xmx1459m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2918m</code>	<code>-Xmx2918m</code>
<code>mapreduce.map.memory.mb</code>	1824	1824
<code>mapreduce.reduce.memory.mb</code>	3648	3648
<code>yarn.app.mapreduce.am.resource.mb</code>	3648	3648
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	116736	29184
<code>yarn.nodemanager.resource.memory-mb</code>	116736	29184



## c5ad.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

## c5d インスタンス

## c5d.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

## c5d.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

## c5d.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## c5d.9xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	65536	32768
<code>yarn.nodemanager.resource.memory-mb</code>	65536	32768

## c5d.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1502m</code>	<code>-Xmx1502m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3004m</code>	<code>-Xmx3004m</code>
<code>mapreduce.map.memory.mb</code>	1877	1877
<code>mapreduce.reduce.memory.mb</code>	3754	3754
<code>yarn.app.mapreduce.am.resource.mb</code>	3754	3754
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	90112	30048
<code>yarn.nodemanager.resource.memory-mb</code>	90112	30048

## c5d.18xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1547m</code>	<code>-Xmx1547m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3094m</code>	<code>-Xmx3094m</code>
<code>mapreduce.map.memory.mb</code>	1934	1934
<code>mapreduce.reduce.memory.mb</code>	3868	3868
<code>yarn.app.mapreduce.am.resource.mb</code>	3868	3868
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	139264	30960
<code>yarn.nodemanager.resource.memory-mb</code>	139264	30960

## c5d.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1570m</code>	<code>-Xmx1570m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3140m</code>	<code>-Xmx3140m</code>
<code>mapreduce.map.memory.mb</code>	1963	1963
<code>mapreduce.reduce.memory.mb</code>	3926	3926
<code>yarn.app.mapreduce.am.resource.mb</code>	3926	3926

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	188416	31376
<code>yarn.nodemanager.resource.memory-mb</code>	188416	31376

## c5n インスタンス

### c5n.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1613m</code>	<code>-Xmx1613m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3226m</code>	<code>-Xmx3226m</code>
<code>mapreduce.map.memory.mb</code>	2016	2016
<code>mapreduce.reduce.memory.mb</code>	4032	4032
<code>yarn.app.mapreduce.am.resource.mb</code>	4032	4032
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	8064	4032
<code>yarn.nodemanager.resource.memory-mb</code>	8064	4032

## c5n.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1613m	-Xmx1613m
mapreduce.java.opts	-Xmx3226m	-Xmx3226m
mapreduce.map.memory.mb	2016	2016
mapreduce.reduce.memory.mb	4032	4032
yarn.app.mapreduce.am.resource.mb	4032	4032
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	16128	8064
yarn.nodemanager.resource.memory-mb	16128	8064

## c5n.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1741m	-Xmx1741m
mapreduce.java.opts	-Xmx3482m	-Xmx3482m
mapreduce.map.memory.mb	2176	2176
mapreduce.reduce.memory.mb	4352	4352
yarn.app.mapreduce.am.resource.mb	4352	4352

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	34816	17408
<code>yarn.nodemanager.resource.memory-mb</code>	34816	17408

## c5n.9xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2002m</code>	<code>-Xmx2002m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4004m</code>	<code>-Xmx4004m</code>
<code>mapreduce.map.memory.mb</code>	2503	2503
<code>mapreduce.reduce.memory.mb</code>	5006	5006
<code>yarn.app.mapreduce.am.resource.mb</code>	5006	5006
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	90112	30040
<code>yarn.nodemanager.resource.memory-mb</code>	90112	30040



## c5n.18xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2094m	-Xmx2094m
mapreduce.java.opts	-Xmx4188m	-Xmx4188m
mapreduce.map.memory.mb	2617	2617
mapreduce.reduce.memory.mb	5234	5234
yarn.app.mapreduce.am.resource.mb	5234	5234
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

## c6a インスタンス

## c6a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c6a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c6a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c6a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c6a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1425m</code>	<code>-Xmx1425m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.map.memory.mb</code>	1781	1781
<code>mapreduce.reduce.memory.mb</code>	3562	3562
<code>yarn.app.mapreduce.am.resource.mb</code>	3562	3562
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074

## c6a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c6a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29840
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29840

## c6a.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1510m</code>	<code>-Xmx1510m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3020m</code>	<code>-Xmx3020m</code>
<code>mapreduce.map.memory.mb</code>	1888	1888
<code>mapreduce.reduce.memory.mb</code>	3776	3776
<code>yarn.app.mapreduce.am.resource.mb</code>	3776	3776
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## c6a.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

## c6g インスタンス

## c6g.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c6g.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856



## c6g.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c6g.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## c6g.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

## c6g.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c6gd インスタンス

## c6gd.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c6gd.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c6gd.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c6gd.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c6gd.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1425m</code>	<code>-Xmx1425m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.map.memory.mb</code>	1781	1781
<code>mapreduce.reduce.memory.mb</code>	3562	3562
<code>yarn.app.mapreduce.am.resource.mb</code>	3562	3562
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074

## c6gd.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c6gn インスタンス

## c6gn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c6gn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856



## c6gn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c6gn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c6gn.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1425m</code>	<code>-Xmx1425m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.map.memory.mb</code>	1781	1781
<code>mapreduce.reduce.memory.mb</code>	3562	3562
<code>yarn.app.mapreduce.am.resource.mb</code>	3562	3562
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074

## c6gn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c6i インスタンス

## c6i.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c6i.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c6i.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c6i.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## c6i.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

## c6i.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c6i.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29840
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29840

## c6i.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1510m</code>	<code>-Xmx1510m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3020m</code>	<code>-Xmx3020m</code>
<code>mapreduce.map.memory.mb</code>	1888	1888
<code>mapreduce.reduce.memory.mb</code>	3776	3776
<code>yarn.app.mapreduce.am.resource.mb</code>	3776	3776
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208



## c6id インスタンス

## c6id.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c6id.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c6id.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c6id.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## c6id.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074

## c6id.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1459m</code>	<code>-Xmx1459m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2918m</code>	<code>-Xmx2918m</code>
<code>mapreduce.map.memory.mb</code>	1824	1824
<code>mapreduce.reduce.memory.mb</code>	3648	3648
<code>yarn.app.mapreduce.am.resource.mb</code>	3648	3648
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	116736	29184
<code>yarn.nodemanager.resource.memory-mb</code>	116736	29184

## c6id.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

## c6id.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## c6in インスタンス

### c6in.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1126m</code>	<code>-Xmx1126m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2252m</code>	<code>-Xmx2252m</code>
<code>mapreduce.map.memory.mb</code>	1408	1408
<code>mapreduce.reduce.memory.mb</code>	2816	2816
<code>yarn.app.mapreduce.am.resource.mb</code>	2816	2816
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	5632	2816
<code>yarn.nodemanager.resource.memory-mb</code>	5632	2816

## c6in.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c6in.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23424	11712
<code>yarn.nodemanager.resource.memory-mb</code>	23424	11712

## c6in.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1357m</code>	<code>-Xmx1357m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2714m</code>	<code>-Xmx2714m</code>
<code>mapreduce.map.memory.mb</code>	1696	1696
<code>mapreduce.reduce.memory.mb</code>	3392	3392
<code>yarn.app.mapreduce.am.resource.mb</code>	3392	3392
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136



## c6in.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

## c6in.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c6in.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

## c6in.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## c7a インスタンス

## c7a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c7a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c7a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c7a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c7a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1425m</code>	<code>-Xmx1425m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.map.memory.mb</code>	1781	1781
<code>mapreduce.reduce.memory.mb</code>	3562	3562
<code>yarn.app.mapreduce.am.resource.mb</code>	3562	3562
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074

## c7a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c7a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29840
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29840

## c7a.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1510m</code>	<code>-Xmx1510m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3020m</code>	<code>-Xmx3020m</code>
<code>mapreduce.map.memory.mb</code>	1888	1888
<code>mapreduce.reduce.memory.mb</code>	3776	3776
<code>yarn.app.mapreduce.am.resource.mb</code>	3776	3776
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208



## c7a.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

## c7g インスタンス

## c7g.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c7g.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c7g.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c7g.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c7g.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1425m</code>	<code>-Xmx1425m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.map.memory.mb</code>	1781	1781
<code>mapreduce.reduce.memory.mb</code>	3562	3562
<code>yarn.app.mapreduce.am.resource.mb</code>	3562	3562
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074

## c7g.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c7gd インスタンス

## c7gd.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c7gd.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c7gd.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c7gd.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c7gd.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1425m</code>	<code>-Xmx1425m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.map.memory.mb</code>	1781	1781
<code>mapreduce.reduce.memory.mb</code>	3562	3562
<code>yarn.app.mapreduce.am.resource.mb</code>	3562	3562
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074



## c7gd.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c7gn インスタンス

## c7gn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c7gn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c7gn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c7gn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c7gn.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1425m</code>	<code>-Xmx1425m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.map.memory.mb</code>	1781	1781
<code>mapreduce.reduce.memory.mb</code>	3562	3562
<code>yarn.app.mapreduce.am.resource.mb</code>	3562	3562
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074

## c7gn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c7i インスタンス

## c7i.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

## c7i.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## c7i.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## c7i.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## c7i.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1425m</code>	<code>-Xmx1425m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.map.memory.mb</code>	1781	1781
<code>mapreduce.reduce.memory.mb</code>	3562	3562
<code>yarn.app.mapreduce.am.resource.mb</code>	3562	3562
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	32074
<code>yarn.nodemanager.resource.memory-mb</code>	85504	32074



## c7i.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## c7i.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29840
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29840

## c7i.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1527m</code>	<code>-Xmx1527m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3054m</code>	<code>-Xmx3054m</code>
<code>mapreduce.map.memory.mb</code>	1909	1909
<code>mapreduce.reduce.memory.mb</code>	3818	3818
<code>yarn.app.mapreduce.am.resource.mb</code>	3818	3818
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30608
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30608

## d2 インスタンス

## d2.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## d2.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## d2.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## d2.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2417m	-Xmx2417m
mapreduce.java.opts	-Xmx4834m	-Xmx4834m
mapreduce.map.memory.mb	3021	3021
mapreduce.reduce.memory.mb	6042	6042
yarn.app.mapreduce.am.resource.mb	6042	6042
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30194
yarn.nodemanager.resource.memory-mb	241664	30194

## d3 インスタンス

## d3.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## d3.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## d3.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## d3.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## d3en インスタンス

### d3en.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2342m</code>	<code>-Xmx2342m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4684m</code>	<code>-Xmx4684m</code>
<code>mapreduce.map.memory.mb</code>	2928	2928
<code>mapreduce.reduce.memory.mb</code>	5856	5856
<code>yarn.app.mapreduce.am.resource.mb</code>	5856	5856
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11712	5856
<code>yarn.nodemanager.resource.memory-mb</code>	11712	5856



## d3en.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## d3en.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## d3en.6xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2850m</code>	<code>-Xmx2850m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx5700m</code>	<code>-Xmx5700m</code>
<code>mapreduce.map.memory.mb</code>	3563	3563
<code>mapreduce.reduce.memory.mb</code>	7126	7126
<code>yarn.app.mapreduce.am.resource.mb</code>	7126	7126
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	28496
<code>yarn.nodemanager.resource.memory-mb</code>	85504	28496

## d3en.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## d3en.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29880
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29880

## g3 インスタンス

### g3.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx5837m</code>	<code>-Xmx5837m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx11674m</code>	<code>-Xmx11674m</code>
<code>mapreduce.map.memory.mb</code>	7296	7296
<code>mapreduce.reduce.memory.mb</code>	14592	14592
<code>yarn.app.mapreduce.am.resource.mb</code>	14592	14592
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	116736	29184
<code>yarn.nodemanager.resource.memory-mb</code>	116736	29184

## g3.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## g3.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

### g3s インスタンス

#### g3s.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx4685m</code>	<code>-Xmx4685m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx9370m</code>	<code>-Xmx9370m</code>
<code>mapreduce.map.memory.mb</code>	5856	5856
<code>mapreduce.reduce.memory.mb</code>	11712	11712
<code>yarn.app.mapreduce.am.resource.mb</code>	11712	11712
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23424	11712
<code>yarn.nodemanager.resource.memory-mb</code>	23424	11712

## g4dn インスタンス

## g4dn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

## g4dn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## g4dn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672



## g4dn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

## g4dn.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

## g4dn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

## g5 インスタンス

## g5.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## g5.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## g5.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## g5.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## g5.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29880
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29880

## g5.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3021m</code>	<code>-Xmx3021m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6042m</code>	<code>-Xmx6042m</code>
<code>mapreduce.map.memory.mb</code>	3776	3776
<code>mapreduce.reduce.memory.mb</code>	7552	7552
<code>yarn.app.mapreduce.am.resource.mb</code>	7552	7552
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## g5.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

## g5.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	741376	30952
<code>yarn.nodemanager.resource.memory-mb</code>	741376	30952

## h1 インスタンス

### h1.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2458m</code>	<code>-Xmx2458m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4916m</code>	<code>-Xmx4916m</code>
<code>mapreduce.map.memory.mb</code>	3072	3072
<code>mapreduce.reduce.memory.mb</code>	6144	6144
<code>yarn.app.mapreduce.am.resource.mb</code>	6144	6144
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	24576	12288
<code>yarn.nodemanager.resource.memory-mb</code>	24576	12288



## h1.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## h1.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	122880	30720
<code>yarn.nodemanager.resource.memory-mb</code>	122880	30720

## h1.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3174m</code>	<code>-Xmx3174m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6348m</code>	<code>-Xmx6348m</code>
<code>mapreduce.map.memory.mb</code>	3968	3968
<code>mapreduce.reduce.memory.mb</code>	7936	7936
<code>yarn.app.mapreduce.am.resource.mb</code>	7936	7936
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	253952	31744
<code>yarn.nodemanager.resource.memory-mb</code>	253952	31744

## i2 インスタンス

## i2.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## i2.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## i2.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## i2.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## i3 インスタンス

## i3.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## i3.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## i3.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## i3.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## i3.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720



## i3en インスタンス

## i3en.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## i3en.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## i3en.3xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

## i3en.6xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

## i3en.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	385024	32100
<code>yarn.nodemanager.resource.memory-mb</code>	385024	32100

## i3en.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6486m</code>	<code>-Xmx6486m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12972m</code>	<code>-Xmx12972m</code>
<code>mapreduce.map.memory.mb</code>	8107	8107
<code>mapreduce.reduce.memory.mb</code>	16214	16214
<code>yarn.app.mapreduce.am.resource.mb</code>	16214	16214
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	778240	32396
<code>yarn.nodemanager.resource.memory-mb</code>	778240	32396

## i4g インスタンス

## i4g.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## i4g.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## i4g.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## i4g.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## i4g.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## i4i インスタンス

### i4i.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx4685m</code>	<code>-Xmx4685m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx9370m</code>	<code>-Xmx9370m</code>
<code>mapreduce.map.memory.mb</code>	5856	5856
<code>mapreduce.reduce.memory.mb</code>	11712	11712
<code>yarn.app.mapreduce.am.resource.mb</code>	11712	11712
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23424	11712
<code>yarn.nodemanager.resource.memory-mb</code>	23424	11712



## i4i.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## i4i.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## i4i.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## i4i.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

## i4i.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## i4i.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6178m</code>	<code>-Xmx6178m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12356m</code>	<code>-Xmx12356m</code>
<code>mapreduce.map.memory.mb</code>	7723	7723
<code>mapreduce.reduce.memory.mb</code>	15446	15446
<code>yarn.app.mapreduce.am.resource.mb</code>	15446	15446
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	741376	30860
<code>yarn.nodemanager.resource.memory-mb</code>	741376	30860

## i4i.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

## im4gn インスタンス

## im4gn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## im4gn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## im4gn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## im4gn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## im4gn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208



## is4gen インスタンス

## is4gen.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3514m	-Xmx3514m
mapreduce.java.opts	-Xmx7028m	-Xmx7028m
mapreduce.map.memory.mb	4393	4393
mapreduce.reduce.memory.mb	8786	8786
yarn.app.mapreduce.am.resource.mb	8786	8786
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	17572.12	8786.06
yarn.nodemanager.resource.memory-mb	17572.12	8786.06

## is4gen.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3866m	-Xmx3866m
mapreduce.java.opts	-Xmx7732m	-Xmx7732m
mapreduce.map.memory.mb	4832	4832
mapreduce.reduce.memory.mb	9664	9664

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	9664	9664
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

## is4gen.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4275m	-Xmx4275m
mapreduce.java.opts	-Xmx8550m	-Xmx8550m
mapreduce.map.memory.mb	5344	5344
mapreduce.reduce.memory.mb	10688	10688
yarn.app.mapreduce.am.resource.mb	10688	10688
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	85504	32064
yarn.nodemanager.resource.memory-mb	85504	32064

## is4gen.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4480m	-Xmx4480m
mapreduce.java.opts	-Xmx8960m	-Xmx8960m
mapreduce.map.memory.mb	5600	5600
mapreduce.reduce.memory.mb	11200	11200
yarn.app.mapreduce.am.resource.mb	11200	11200
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	22400
yarn.nodemanager.resource.memory-mb	179200	22400

## m1 インスタンス

## m1.small

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx288m	-Xmx288m
mapreduce.java.opts	-Xmx288m	-Xmx288m
mapreduce.map.memory.mb	512	512
mapreduce.reduce.memory.mb	512	512

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	512	512
yarn.scheduler.minimum-allocation-mb	256	256
yarn.scheduler.maximum-allocation-mb	512	512
yarn.nodemanager.resource.memory-mb	1024	512

## m1.medium

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx768m	-Xmx768m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	1024	1024
yarn.app.mapreduce.am.resource.mb	1024	1024
yarn.scheduler.minimum-allocation-mb	256	256
yarn.scheduler.maximum-allocation-mb	2048	1024
yarn.nodemanager.resource.memory-mb	2048	1024

## m1.large

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1024m	-Xmx1024m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	1536	1536
yarn.app.mapreduce.am.resource.mb	1536	1536
yarn.scheduler.minimum-allocation-mb	256	256
yarn.scheduler.maximum-allocation-mb	3072	2560
yarn.nodemanager.resource.memory-mb	5120	2560

## m1.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	256	256
<code>yarn.scheduler.maximum-allocation-mb</code>	8192	6144
<code>yarn.nodemanager.resource.memory-mb</code>	12288	6144

## m2 インスタンス

### m2.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx864m</code>	<code>-Xmx864m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx1536m</code>	<code>-Xmx1536m</code>
<code>mapreduce.map.memory.mb</code>	1024	1024
<code>mapreduce.reduce.memory.mb</code>	2048	2048
<code>yarn.app.mapreduce.am.resource.mb</code>	2048	2048
<code>yarn.scheduler.minimum-allocation-mb</code>	256	256
<code>yarn.scheduler.maximum-allocation-mb</code>	7168	7168
<code>yarn.nodemanager.resource.memory-mb</code>	14336	7168

## m2.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560
yarn.scheduler.minimum-allocation-mb	256	256
yarn.scheduler.maximum-allocation-mb	8192	15360
yarn.nodemanager.resource.memory-mb	30720	15360

## m2.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	256	256
<code>yarn.scheduler.maximum-allocation-mb</code>	8192	30720
<code>yarn.nodemanager.resource.memory-mb</code>	61440	30720

### m3 インスタンス

#### m3.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1152m</code>	<code>-Xmx1152m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2304m</code>	<code>-Xmx2304m</code>
<code>mapreduce.map.memory.mb</code>	1440	1440
<code>mapreduce.reduce.memory.mb</code>	2880	2880
<code>yarn.app.mapreduce.am.resource.mb</code>	2880	2880
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11520	5760
<code>yarn.nodemanager.resource.memory-mb</code>	11520	5760



## m3.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1152m</code>	<code>-Xmx1152m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2304m</code>	<code>-Xmx2304m</code>
<code>mapreduce.map.memory.mb</code>	1440	1440
<code>mapreduce.reduce.memory.mb</code>	2880	2880
<code>yarn.app.mapreduce.am.resource.mb</code>	2880	2880
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23040	11520
<code>yarn.nodemanager.resource.memory-mb</code>	23040	11520

## m4 インスタンス

## m4.large

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1229m</code>	<code>-Xmx1229m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx2458m</code>	<code>-Xmx2458m</code>
<code>mapreduce.map.memory.mb</code>	1536	1536
<code>mapreduce.reduce.memory.mb</code>	3072	3072

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

## m4.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

## m4.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## m4.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1434m	-Xmx1434m
mapreduce.java.opts	-Xmx2868m	-Xmx2868m
mapreduce.map.memory.mb	1792	1792
mapreduce.reduce.memory.mb	3584	3584
yarn.app.mapreduce.am.resource.mb	3584	3584

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	57344	28672
<code>yarn.nodemanager.resource.memory-mb</code>	57344	28672

## m4.10xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx1557m</code>	<code>-Xmx1557m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx3114m</code>	<code>-Xmx3114m</code>
<code>mapreduce.map.memory.mb</code>	1946	1946
<code>mapreduce.reduce.memory.mb</code>	3892	3892
<code>yarn.app.mapreduce.am.resource.mb</code>	3892	3892
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	155648	31104
<code>yarn.nodemanager.resource.memory-mb</code>	155648	31104

## m4.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx1587m	-Xmx1587m
mapreduce.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.map.memory.mb	1984	1984
mapreduce.reduce.memory.mb	3968	3968
yarn.app.mapreduce.am.resource.mb	3968	3968
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

## m5 インスタンス

## m5.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

## m5.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## m5.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## m5.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

## m5.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416



## m5.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

## m5.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	385024	32056
<code>yarn.nodemanager.resource.memory-mb</code>	385024	32056

## m5a インスタンス

### m5a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2458m</code>	<code>-Xmx2458m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4916m</code>	<code>-Xmx4916m</code>
<code>mapreduce.map.memory.mb</code>	3072	3072
<code>mapreduce.reduce.memory.mb</code>	6144	6144
<code>yarn.app.mapreduce.am.resource.mb</code>	6144	6144
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	12288	6144
<code>yarn.nodemanager.resource.memory-mb</code>	12288	6144

## m5a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## m5a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## m5a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

## m5a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

## m5a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

## m5a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

## m5ad インスタンス

## m5ad.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m5ad.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m5ad.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136



## m5ad.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m5ad.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29880
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29880

## m5ad.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3021m</code>	<code>-Xmx3021m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6042m</code>	<code>-Xmx6042m</code>
<code>mapreduce.map.memory.mb</code>	3776	3776
<code>mapreduce.reduce.memory.mb</code>	7552	7552
<code>yarn.app.mapreduce.am.resource.mb</code>	7552	7552
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## m5ad.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

## m5d インスタンス

## m5d.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

## m5d.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## m5d.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## m5d.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	122880	30720
<code>yarn.nodemanager.resource.memory-mb</code>	122880	30720

## m5d.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3140m</code>	<code>-Xmx3140m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6280m</code>	<code>-Xmx6280m</code>
<code>mapreduce.map.memory.mb</code>	3925	3925
<code>mapreduce.reduce.memory.mb</code>	7850	7850
<code>yarn.app.mapreduce.am.resource.mb</code>	7850	7850
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	188416	31416
<code>yarn.nodemanager.resource.memory-mb</code>	188416	31416

## m5d.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

## m5d.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	385024	32056
<code>yarn.nodemanager.resource.memory-mb</code>	385024	32056

## m5dn インスタンス

### m5dn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2342m</code>	<code>-Xmx2342m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4684m</code>	<code>-Xmx4684m</code>
<code>mapreduce.map.memory.mb</code>	2928	2928
<code>mapreduce.reduce.memory.mb</code>	5856	5856
<code>yarn.app.mapreduce.am.resource.mb</code>	5856	5856
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11712	5856
<code>yarn.nodemanager.resource.memory-mb</code>	11712	5856



## m5dn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m5dn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m5dn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m5dn.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

## m5dn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## m5dn.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3055m</code>	<code>-Xmx3055m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6110m</code>	<code>-Xmx6110m</code>
<code>mapreduce.map.memory.mb</code>	3819	3819
<code>mapreduce.reduce.memory.mb</code>	7638	7638
<code>yarn.app.mapreduce.am.resource.mb</code>	7638	7638
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30520
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30520

## m5n インスタンス

## m5n.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m5n.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m5n.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m5n.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m5n.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29880
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29880

## m5n.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3021m</code>	<code>-Xmx3021m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6042m</code>	<code>-Xmx6042m</code>
<code>mapreduce.map.memory.mb</code>	3776	3776
<code>mapreduce.reduce.memory.mb</code>	7552	7552
<code>yarn.app.mapreduce.am.resource.mb</code>	7552	7552
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208



## m5n.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

## m5zn インスタンス

## m5zn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2880	2880
mapreduce.reduce.memory.mb	5760	5760

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5760	5760
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

## m5zn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m5zn.3xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2577m	-Xmx2577m
mapreduce.java.opts	-Xmx5154m	-Xmx5154m
mapreduce.map.memory.mb	3221	3221
mapreduce.reduce.memory.mb	6442	6442
yarn.app.mapreduce.am.resource.mb	6442	6442
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

## m5zn.6xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	85504	28496
<code>yarn.nodemanager.resource.memory-mb</code>	85504	28496

## m5zn.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2986m</code>	<code>-Xmx2986m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx5972m</code>	<code>-Xmx5972m</code>
<code>mapreduce.map.memory.mb</code>	3733	3733
<code>mapreduce.reduce.memory.mb</code>	7466	7466
<code>yarn.app.mapreduce.am.resource.mb</code>	7466	7466
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29880
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29880

## m6a インスタンス

## m6a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m6a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m6a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m6a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m6a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

## m6a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208



## m6a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

## m6a.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## m6a.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3089m</code>	<code>-Xmx3089m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6178m</code>	<code>-Xmx6178m</code>
<code>mapreduce.map.memory.mb</code>	3861	3861
<code>mapreduce.reduce.memory.mb</code>	7722	7722
<code>yarn.app.mapreduce.am.resource.mb</code>	7722	7722
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	741376	30952
<code>yarn.nodemanager.resource.memory-mb</code>	741376	30952

## m6g インスタンス

## m6g.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m6g.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m6g.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m6g.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m6g.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	181248	30208
<code>yarn.nodemanager.resource.memory-mb</code>	181248	30208

## m6g.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3021m</code>	<code>-Xmx3021m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6042m</code>	<code>-Xmx6042m</code>
<code>mapreduce.map.memory.mb</code>	3776	3776
<code>mapreduce.reduce.memory.mb</code>	7552	7552
<code>yarn.app.mapreduce.am.resource.mb</code>	7552	7552
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## m6gd インスタンス

## m6gd.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m6gd.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m6gd.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136



## m6gd.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m6gd.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	181248	30208
<code>yarn.nodemanager.resource.memory-mb</code>	181248	30208

## m6gd.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3021m</code>	<code>-Xmx3021m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6042m</code>	<code>-Xmx6042m</code>
<code>mapreduce.map.memory.mb</code>	3776	3776
<code>mapreduce.reduce.memory.mb</code>	7552	7552
<code>yarn.app.mapreduce.am.resource.mb</code>	7552	7552
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## m6i インスタンス

## m6i.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m6i.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m6i.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m6i.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m6i.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

## m6i.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## m6i.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

## m6i.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## m6id インスタンス

### m6id.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2342m</code>	<code>-Xmx2342m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4684m</code>	<code>-Xmx4684m</code>
<code>mapreduce.map.memory.mb</code>	2928	2928
<code>mapreduce.reduce.memory.mb</code>	5856	5856
<code>yarn.app.mapreduce.am.resource.mb</code>	5856	5856
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11712	5856
<code>yarn.nodemanager.resource.memory-mb</code>	11712	5856



## m6id.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m6id.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## m6id.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2918m</code>	<code>-Xmx2918m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx5836m</code>	<code>-Xmx5836m</code>
<code>mapreduce.map.memory.mb</code>	3648	3648
<code>mapreduce.reduce.memory.mb</code>	7296	7296
<code>yarn.app.mapreduce.am.resource.mb</code>	7296	7296
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	116736	29184
<code>yarn.nodemanager.resource.memory-mb</code>	116736	29184

## m6id.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

## m6id.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## m6id.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

## m6id.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## m6idn インスタンス

## m6idn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m6idn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m6idn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m6idn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	116736	29184
<code>yarn.nodemanager.resource.memory-mb</code>	116736	29184

## m6idn.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2986m</code>	<code>-Xmx2986m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx5972m</code>	<code>-Xmx5972m</code>
<code>mapreduce.map.memory.mb</code>	3733	3733
<code>mapreduce.reduce.memory.mb</code>	7466	7466
<code>yarn.app.mapreduce.am.resource.mb</code>	7466	7466
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	179200	29880
<code>yarn.nodemanager.resource.memory-mb</code>	179200	29880



## m6idn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## m6idn.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

## m6idn.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## m6in インスタンス

## m6in.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m6in.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m6in.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m6in.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m6in.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

## m6in.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## m6in.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

## m6in.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## m7a インスタンス

### m7a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2342m</code>	<code>-Xmx2342m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4684m</code>	<code>-Xmx4684m</code>
<code>mapreduce.map.memory.mb</code>	2928	2928
<code>mapreduce.reduce.memory.mb</code>	5856	5856
<code>yarn.app.mapreduce.am.resource.mb</code>	5856	5856
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11712	5856
<code>yarn.nodemanager.resource.memory-mb</code>	11712	5856



## m7a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m7a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m7a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m7a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

## m7a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## m7a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3055m</code>	<code>-Xmx3055m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6110m</code>	<code>-Xmx6110m</code>
<code>mapreduce.map.memory.mb</code>	3819	3819
<code>mapreduce.reduce.memory.mb</code>	7638	7638
<code>yarn.app.mapreduce.am.resource.mb</code>	7638	7638
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30520
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30520

## m7a.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## m7a.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	741376	30952
<code>yarn.nodemanager.resource.memory-mb</code>	741376	30952

## m7g インスタンス

### m7g.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2342m</code>	<code>-Xmx2342m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4684m</code>	<code>-Xmx4684m</code>
<code>mapreduce.map.memory.mb</code>	2928	2928
<code>mapreduce.reduce.memory.mb</code>	5856	5856
<code>yarn.app.mapreduce.am.resource.mb</code>	5856	5856
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11712	5856
<code>yarn.nodemanager.resource.memory-mb</code>	11712	5856

## m7g.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m7g.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m7g.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184



## m7g.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

## m7g.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## m7gd インスタンス

### m7gd.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2342m</code>	<code>-Xmx2342m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4684m</code>	<code>-Xmx4684m</code>
<code>mapreduce.map.memory.mb</code>	2928	2928
<code>mapreduce.reduce.memory.mb</code>	5856	5856
<code>yarn.app.mapreduce.am.resource.mb</code>	5856	5856
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11712	5856
<code>yarn.nodemanager.resource.memory-mb</code>	11712	5856

## m7gd.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m7gd.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m7gd.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## m7gd.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

## m7gd.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## m7i インスタンス

### m7i.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx2342m</code>	<code>-Xmx2342m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx4684m</code>	<code>-Xmx4684m</code>
<code>mapreduce.map.memory.mb</code>	2928	2928
<code>mapreduce.reduce.memory.mb</code>	5856	5856
<code>yarn.app.mapreduce.am.resource.mb</code>	5856	5856
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	11712	5856
<code>yarn.nodemanager.resource.memory-mb</code>	11712	5856

## m7i.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m7i.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m7i.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184



## m7i.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

## m7i.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## m7i.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx3055m</code>	<code>-Xmx3055m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx6110m</code>	<code>-Xmx6110m</code>
<code>mapreduce.map.memory.mb</code>	3819	3819
<code>mapreduce.reduce.memory.mb</code>	7638	7638
<code>yarn.app.mapreduce.am.resource.mb</code>	7638	7638
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30520
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30520

## m7i.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

## m7i-flex インスタンス

## m7i-flex.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

## m7i-flex.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## m7i-flex.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## m7i-flex.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	116736	29184
<code>yarn.nodemanager.resource.memory-mb</code>	116736	29184

## p2 インスタンス

### p2.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx10854m</code>	<code>-Xmx10854m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx21708m</code>	<code>-Xmx21708m</code>
<code>mapreduce.map.memory.mb</code>	13568	13568
<code>mapreduce.reduce.memory.mb</code>	27136	27136
<code>yarn.app.mapreduce.am.resource.mb</code>	27136	27136
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136

## p2.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## p2.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx9267m	-Xmx9267m
mapreduce.java.opts	-Xmx18534m	-Xmx18534m
mapreduce.map.memory.mb	11584	11584
mapreduce.reduce.memory.mb	23168	23168
yarn.app.mapreduce.am.resource.mb	23168	23168

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	741376	23168
<code>yarn.nodemanager.resource.memory-mb</code>	741376	23168

### p3 インスタンス

#### p3.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx5427m</code>	<code>-Xmx5427m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx10854m</code>	<code>-Xmx10854m</code>
<code>mapreduce.map.memory.mb</code>	6784	6784
<code>mapreduce.reduce.memory.mb</code>	13568	13568
<code>yarn.app.mapreduce.am.resource.mb</code>	13568	13568
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	54272	27136
<code>yarn.nodemanager.resource.memory-mb</code>	54272	27136



## p3.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## p3.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## p5 インスタンス

### p5.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx8294m</code>	<code>-Xmx8294m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx16588m</code>	<code>-Xmx16588m</code>
<code>mapreduce.map.memory.mb</code>	10368	10368
<code>mapreduce.reduce.memory.mb</code>	20736	20736
<code>yarn.app.mapreduce.am.resource.mb</code>	20736	20736
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	1990656	20736
<code>yarn.nodemanager.resource.memory-mb</code>	1990656	20736

## r3 インスタンス

## r3.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r3.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r3.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r3.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r4 インスタンス

## r4.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r4.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r4.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r4.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r4.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720



## r5 インスタンス

## r5.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## r5.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## r5.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

## r5.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

## r5.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	385024	32100
<code>yarn.nodemanager.resource.memory-mb</code>	385024	32100

## r5.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6451m</code>	<code>-Xmx6451m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12902m</code>	<code>-Xmx12902m</code>
<code>mapreduce.map.memory.mb</code>	8064	8064
<code>mapreduce.reduce.memory.mb</code>	16128	16128
<code>yarn.app.mapreduce.am.resource.mb</code>	16128	16128
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	516096	32256
<code>yarn.nodemanager.resource.memory-mb</code>	516096	32256

## r5.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

## r5a インスタンス

## r5a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## r5a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## r5a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

## r5a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	253952	31744
<code>yarn.nodemanager.resource.memory-mb</code>	253952	31744

## r5a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6417m</code>	<code>-Xmx6417m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12834m</code>	<code>-Xmx12834m</code>
<code>mapreduce.map.memory.mb</code>	8021	8021
<code>mapreduce.reduce.memory.mb</code>	16042	16042
<code>yarn.app.mapreduce.am.resource.mb</code>	16042	16042
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	385024	32100
<code>yarn.nodemanager.resource.memory-mb</code>	385024	32100



## r5a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

## r5a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	778240	32396
<code>yarn.nodemanager.resource.memory-mb</code>	778240	32396

## r5ad インスタンス

### r5ad.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx4685m</code>	<code>-Xmx4685m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx9370m</code>	<code>-Xmx9370m</code>
<code>mapreduce.map.memory.mb</code>	5856	5856
<code>mapreduce.reduce.memory.mb</code>	11712	11712
<code>yarn.app.mapreduce.am.resource.mb</code>	11712	11712
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23424	11712
<code>yarn.nodemanager.resource.memory-mb</code>	23424	11712

## r5ad.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r5ad.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r5ad.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r5ad.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

## r5ad.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6246m	-Xmx6246m
mapreduce.java.opts	-Xmx12492m	-Xmx12492m
mapreduce.map.memory.mb	7808	7808
mapreduce.reduce.memory.mb	15616	15616
yarn.app.mapreduce.am.resource.mb	15616	15616

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	499712	31232
yarn.nodemanager.resource.memory-mb	499712	31232

## r5ad.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r5b インスタンス

## r5b.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r5b.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r5b.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184



## r5b.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r5b.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r5b.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6144m</code>	<code>-Xmx6144m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12288m</code>	<code>-Xmx12288m</code>
<code>mapreduce.map.memory.mb</code>	7680	7680
<code>mapreduce.reduce.memory.mb</code>	15360	15360
<code>yarn.app.mapreduce.am.resource.mb</code>	15360	15360
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## r5b.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r5d インスタンス

## r5d.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

## r5d.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## r5d.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

## r5d.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	253952	31744
<code>yarn.nodemanager.resource.memory-mb</code>	253952	31744

## r5d.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6417m</code>	<code>-Xmx6417m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12834m</code>	<code>-Xmx12834m</code>
<code>mapreduce.map.memory.mb</code>	8021	8021
<code>mapreduce.reduce.memory.mb</code>	16042	16042
<code>yarn.app.mapreduce.am.resource.mb</code>	16042	16042
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	385024	32100
<code>yarn.nodemanager.resource.memory-mb</code>	385024	32100

## r5d.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

## r5d.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	778240	32396
<code>yarn.nodemanager.resource.memory-mb</code>	778240	32396

## r5dn インスタンス

### r5dn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx4685m</code>	<code>-Xmx4685m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx9370m</code>	<code>-Xmx9370m</code>
<code>mapreduce.map.memory.mb</code>	5856	5856
<code>mapreduce.reduce.memory.mb</code>	11712	11712
<code>yarn.app.mapreduce.am.resource.mb</code>	11712	11712
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23424	11712
<code>yarn.nodemanager.resource.memory-mb</code>	23424	11712



## r5dn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r5dn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r5dn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r5dn.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

## r5dn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## r5dn.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r5n インスタンス

## r5n.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r5n.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r5n.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r5n.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r5n.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

## r5n.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720



## r5n.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r6a インスタンス

## r6a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r6a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r6a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r6a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## r6a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6110m</code>	<code>-Xmx6110m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12220m</code>	<code>-Xmx12220m</code>
<code>mapreduce.map.memory.mb</code>	7637	7637
<code>mapreduce.reduce.memory.mb</code>	15274	15274
<code>yarn.app.mapreduce.am.resource.mb</code>	15274	15274
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r6a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## r6a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r6a.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

## r6a.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

## r6g インスタンス

## r6g.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r6g.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136



## r6g.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r6g.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## r6g.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6110m</code>	<code>-Xmx6110m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12220m</code>	<code>-Xmx12220m</code>
<code>mapreduce.map.memory.mb</code>	7637	7637
<code>mapreduce.reduce.memory.mb</code>	15274	15274
<code>yarn.app.mapreduce.am.resource.mb</code>	15274	15274
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r6g.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## r6gd インスタンス

## r6gd.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r6gd.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r6gd.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r6gd.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## r6gd.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6110m</code>	<code>-Xmx6110m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12220m</code>	<code>-Xmx12220m</code>
<code>mapreduce.map.memory.mb</code>	7637	7637
<code>mapreduce.reduce.memory.mb</code>	15274	15274
<code>yarn.app.mapreduce.am.resource.mb</code>	15274	15274
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r6gd.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## r6i インスタンス

## r6i.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r6i.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136



## r6i.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r6i.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## r6i.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6110m</code>	<code>-Xmx6110m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12220m</code>	<code>-Xmx12220m</code>
<code>mapreduce.map.memory.mb</code>	7637	7637
<code>mapreduce.reduce.memory.mb</code>	15274	15274
<code>yarn.app.mapreduce.am.resource.mb</code>	15274	15274
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r6i.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## r6i.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r6i.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

## r6id インスタンス

## r6id.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r6id.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r6id.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r6id.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r6id.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r6id.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6144m</code>	<code>-Xmx6144m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12288m</code>	<code>-Xmx12288m</code>
<code>mapreduce.map.memory.mb</code>	7680	7680
<code>mapreduce.reduce.memory.mb</code>	15360	15360
<code>yarn.app.mapreduce.am.resource.mb</code>	15360	15360
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720



## r6id.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r6id.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	991232	30976
<code>yarn.nodemanager.resource.memory-mb</code>	991232	30976

## r6idn インスタンス

### r6idn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx4685m</code>	<code>-Xmx4685m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx9370m</code>	<code>-Xmx9370m</code>
<code>mapreduce.map.memory.mb</code>	5856	5856
<code>mapreduce.reduce.memory.mb</code>	11712	11712
<code>yarn.app.mapreduce.am.resource.mb</code>	11712	11712
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23424	11712
<code>yarn.nodemanager.resource.memory-mb</code>	23424	11712

## r6idn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r6idn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r6idn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r6idn.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

## r6idn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## r6idn.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6178m</code>	<code>-Xmx6178m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12356m</code>	<code>-Xmx12356m</code>
<code>mapreduce.map.memory.mb</code>	7723	7723
<code>mapreduce.reduce.memory.mb</code>	15446	15446
<code>yarn.app.mapreduce.am.resource.mb</code>	15446	15446
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	741376	30860
<code>yarn.nodemanager.resource.memory-mb</code>	741376	30860

## r6idn.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

## r6in インスタンス

## r6in.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r6in.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136



## r6in.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r6in.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	241664	30208
<code>yarn.nodemanager.resource.memory-mb</code>	241664	30208

## r6in.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6110m</code>	<code>-Xmx6110m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12220m</code>	<code>-Xmx12220m</code>
<code>mapreduce.map.memory.mb</code>	7637	7637
<code>mapreduce.reduce.memory.mb</code>	15274	15274
<code>yarn.app.mapreduce.am.resource.mb</code>	15274	15274
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r6in.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

## r6in.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	741376	30860
<code>yarn.nodemanager.resource.memory-mb</code>	741376	30860

## r6in.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6195m</code>	<code>-Xmx6195m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12390m</code>	<code>-Xmx12390m</code>
<code>mapreduce.map.memory.mb</code>	7744	7744
<code>mapreduce.reduce.memory.mb</code>	15488	15488
<code>yarn.app.mapreduce.am.resource.mb</code>	15488	15488
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	991232	30976
<code>yarn.nodemanager.resource.memory-mb</code>	991232	30976

## r7a インスタンス

## r7a.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r7a.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r7a.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r7a.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r7a.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r7a.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6144m</code>	<code>-Xmx6144m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12288m</code>	<code>-Xmx12288m</code>
<code>mapreduce.map.memory.mb</code>	7680	7680
<code>mapreduce.reduce.memory.mb</code>	15360	15360
<code>yarn.app.mapreduce.am.resource.mb</code>	15360	15360
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720



## r7a.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r7a.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

## r7a.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

## r7g インスタンス

## r7g.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r7g.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r7g.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r7g.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r7g.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r7g.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6144m</code>	<code>-Xmx6144m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12288m</code>	<code>-Xmx12288m</code>
<code>mapreduce.map.memory.mb</code>	7680	7680
<code>mapreduce.reduce.memory.mb</code>	15360	15360
<code>yarn.app.mapreduce.am.resource.mb</code>	15360	15360
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## r7gd インスタンス

## r7gd.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r7gd.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r7gd.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184



## r7gd.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r7gd.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r7gd.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6144m</code>	<code>-Xmx6144m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12288m</code>	<code>-Xmx12288m</code>
<code>mapreduce.map.memory.mb</code>	7680	7680
<code>mapreduce.reduce.memory.mb</code>	15360	15360
<code>yarn.app.mapreduce.am.resource.mb</code>	15360	15360
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## r7i インスタンス

## r7i.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

## r7i.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r7i.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r7i.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r7i.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	366592	30564
<code>yarn.nodemanager.resource.memory-mb</code>	366592	30564

## r7i.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6144m</code>	<code>-Xmx6144m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12288m</code>	<code>-Xmx12288m</code>
<code>mapreduce.map.memory.mb</code>	7680	7680
<code>mapreduce.reduce.memory.mb</code>	15360	15360
<code>yarn.app.mapreduce.am.resource.mb</code>	15360	15360
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## r7i.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

## r7i.48xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	1490944	31124
<code>yarn.nodemanager.resource.memory-mb</code>	1490944	31124

## r7iz インスタンス

### r7iz.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx4685m</code>	<code>-Xmx4685m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx9370m</code>	<code>-Xmx9370m</code>
<code>mapreduce.map.memory.mb</code>	5856	5856
<code>mapreduce.reduce.memory.mb</code>	11712	11712
<code>yarn.app.mapreduce.am.resource.mb</code>	11712	11712
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	23424	11712
<code>yarn.nodemanager.resource.memory-mb</code>	23424	11712



## r7iz.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## r7iz.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## r7iz.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## r7iz.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

## r7iz.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## r7iz.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx6029m</code>	<code>-Xmx6029m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx12058m</code>	<code>-Xmx12058m</code>
<code>mapreduce.map.memory.mb</code>	7536	7536
<code>mapreduce.reduce.memory.mb</code>	15072	15072
<code>yarn.app.mapreduce.am.resource.mb</code>	15072	15072
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	964608	30144
<code>yarn.nodemanager.resource.memory-mb</code>	964608	30144

## x1 インスタンス

## x1.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.java.opts	-Xmx24116m	-Xmx24116m
mapreduce.map.memory.mb	15072	15072
mapreduce.reduce.memory.mb	30144	30144
yarn.app.mapreduce.am.resource.mb	30144	30144
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

## x1.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12109m	-Xmx12109m
mapreduce.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.map.memory.mb	15136	15136
mapreduce.reduce.memory.mb	30272	30272

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	30272	30272
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	1937408	30272
yarn.nodemanager.resource.memory-mb	1937408	30272

## x1e インスタンス

### x1e.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx22682m	-Xmx22682m
mapreduce.java.opts	-Xmx45364m	-Xmx45364m
mapreduce.map.memory.mb	28352	28352
mapreduce.reduce.memory.mb	56704	56704
yarn.app.mapreduce.am.resource.mb	56704	56704
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	113408	0

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.nodemanager.resource.memory-mb	113408	0

## x1e.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx23501m	-Xmx23501m
mapreduce.java.opts	-Xmx47002m	-Xmx47002m
mapreduce.map.memory.mb	29376	29376
mapreduce.reduce.memory.mb	58752	58752
yarn.app.mapreduce.am.resource.mb	58752	58752
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	235008	0
yarn.nodemanager.resource.memory-mb	235008	0

## x1e.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx23910m	-Xmx23910m

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.java.opts	-Xmx47820m	-Xmx47820m
mapreduce.map.memory.mb	29888	29888
mapreduce.reduce.memory.mb	59776	59776
yarn.app.mapreduce.am.resource.mb	59776	59776
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	478208	0
yarn.nodemanager.resource.memory-mb	478208	0

## x1e.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx24115m	-Xmx24115m
mapreduce.java.opts	-Xmx48230m	-Xmx48230m
mapreduce.map.memory.mb	30144	30144
mapreduce.reduce.memory.mb	60288	60288
yarn.app.mapreduce.am.resource.mb	60288	60288
yarn.scheduler.minimum-allocation-mb	32	32



設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.maximum-allocation-mb	964608	0
yarn.nodemanager.resource.memory-mb	964608	0

## x1e.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.java.opts	-Xmx48436m	-Xmx48436m
mapreduce.map.memory.mb	30272	30272
mapreduce.reduce.memory.mb	60544	60544
yarn.app.mapreduce.am.resource.mb	60544	60544
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	1937408	0
yarn.nodemanager.resource.memory-mb	1937408	0

## x1e.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx24269m	-Xmx24269m
mapreduce.java.opts	-Xmx48538m	-Xmx48538m
mapreduce.map.memory.mb	30336	30336
mapreduce.reduce.memory.mb	60672	60672
yarn.app.mapreduce.am.resource.mb	60672	60672
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	3883008	0
yarn.nodemanager.resource.memory-mb	3883008	0

## x2gd インスタンス

## x2gd.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

## x2gd.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.java.opts	-Xmx23348m	-Xmx23348m
mapreduce.map.memory.mb	14592	14592
mapreduce.reduce.memory.mb	29184	29184
yarn.app.mapreduce.am.resource.mb	29184	29184
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

## x2gd.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12083m	-Xmx12083m
mapreduce.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.map.memory.mb	15104	15104
mapreduce.reduce.memory.mb	30208	30208
yarn.app.mapreduce.am.resource.mb	30208	30208
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

## x2gd.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	491520	30720
<code>yarn.nodemanager.resource.memory-mb</code>	491520	30720

## x2gd.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx12356m</code>	<code>-Xmx12356m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx24712m</code>	<code>-Xmx24712m</code>
<code>mapreduce.map.memory.mb</code>	15445	15445
<code>mapreduce.reduce.memory.mb</code>	30890	30890
<code>yarn.app.mapreduce.am.resource.mb</code>	30890	30890
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	741376	30906
<code>yarn.nodemanager.resource.memory-mb</code>	741376	30906

## x2gd.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

## x2idn インスタンス

## x2idn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

## x2idn.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12425m	-Xmx12425m
mapreduce.java.opts	-Xmx24850m	-Xmx24850m
mapreduce.map.memory.mb	15531	15531
mapreduce.reduce.memory.mb	31062	31062
yarn.app.mapreduce.am.resource.mb	31062	31062
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

## x2idn.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx12442m	-Xmx12442m
mapreduce.java.opts	-Xmx24884m	-Xmx24884m
mapreduce.map.memory.mb	15552	15552
mapreduce.reduce.memory.mb	31104	31104
yarn.app.mapreduce.am.resource.mb	31104	31104
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	1990656	31104
yarn.nodemanager.resource.memory-mb	1990656	31104

## x2iedn インスタンス

## x2iedn.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx23347m	-Xmx23347m
mapreduce.java.opts	-Xmx46694m	-Xmx46694m
mapreduce.map.memory.mb	29184	29184
mapreduce.reduce.memory.mb	58368	58368



設定オプション	デフォルト値	HBase がインストールされている場合
yarn.app.mapreduce.am.resource.mb	58368	58368
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	116736	0
yarn.nodemanager.resource.memory-mb	116736	0

## x2iedn.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.java.opts	-Xmx48332m	-Xmx48332m
mapreduce.map.memory.mb	30208	30208
mapreduce.reduce.memory.mb	60416	60416
yarn.app.mapreduce.am.resource.mb	60416	60416
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	241664	0
yarn.nodemanager.resource.memory-mb	241664	0

## x2iedn.4xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.java.opts	-Xmx49152m	-Xmx49152m
mapreduce.map.memory.mb	30720	30720
mapreduce.reduce.memory.mb	61440	61440
yarn.app.mapreduce.am.resource.mb	61440	61440
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	491520	0
yarn.nodemanager.resource.memory-mb	491520	0

## x2iedn.8xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx24781m	-Xmx24781m
mapreduce.java.opts	-Xmx49562m	-Xmx49562m
mapreduce.map.memory.mb	30976	30976
mapreduce.reduce.memory.mb	61952	61952
yarn.app.mapreduce.am.resource.mb	61952	61952

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	991232	0
<code>yarn.nodemanager.resource.memory-mb</code>	991232	0

## x2iedn.16xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx24883m</code>	<code>-Xmx24883m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx49766m</code>	<code>-Xmx49766m</code>
<code>mapreduce.map.memory.mb</code>	31104	31104
<code>mapreduce.reduce.memory.mb</code>	62208	62208
<code>yarn.app.mapreduce.am.resource.mb</code>	62208	62208
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	1990656	0
<code>yarn.nodemanager.resource.memory-mb</code>	1990656	0

## x2iedn.24xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx24918m	-Xmx24918m
mapreduce.java.opts	-Xmx49836m	-Xmx49836m
mapreduce.map.memory.mb	31147	31147
mapreduce.reduce.memory.mb	62294	62294
yarn.app.mapreduce.am.resource.mb	62294	62294
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	2990080	-32
yarn.nodemanager.resource.memory-mb	2990080	-32

## x2iedn.32xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx24934m	-Xmx24934m
mapreduce.java.opts	-Xmx49868m	-Xmx49868m
mapreduce.map.memory.mb	31168	31168
mapreduce.reduce.memory.mb	62336	62336
yarn.app.mapreduce.am.resource.mb	62336	62336

設定オプション	デフォルト値	HBase がインストールされている場合
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	3989504	0
<code>yarn.nodemanager.resource.memory-mb</code>	3989504	0

## z1d インスタンス

### z1d.xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
<code>mapreduce.map.java.opts</code>	<code>-Xmx4915m</code>	<code>-Xmx4915m</code>
<code>mapreduce.java.opts</code>	<code>-Xmx9830m</code>	<code>-Xmx9830m</code>
<code>mapreduce.map.memory.mb</code>	6144	6144
<code>mapreduce.reduce.memory.mb</code>	12288	12288
<code>yarn.app.mapreduce.am.resource.mb</code>	12288	12288
<code>yarn.scheduler.minimum-allocation-mb</code>	32	32
<code>yarn.scheduler.maximum-allocation-mb</code>	24576	12288
<code>yarn.nodemanager.resource.memory-mb</code>	24576	12288

## z1d.2xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

## z1d.3xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018

設定オプション	デフォルト値	HBase がインストールされている場合
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

## z1d.6xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

## z1d.12xlarge

設定オプション	デフォルト値	HBase がインストールされている場合
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	32	32
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

## Hadoop デーモン構成設定

Hadoop デーモン設定は、クラスターノードが使用する EC2 インスタンスタイプによって異なります。以下の表に、EC2 の各インスタンスタイプのデフォルトの設定を示します。

これらの設定をカスタマイズするには、`hadoop-env` 設定分類を使用します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## インスタンスタイプ

- [c1 インスタンス](#)
- [c3 インスタンス](#)
- [c4 インスタンス](#)
- [c5 インスタンス](#)



- [c5a インスタンス](#)
- [c5ad インスタンス](#)
- [c5d インスタンス](#)
- [c5n インスタンス](#)
- [c6a インスタンス](#)
- [c6g インスタンス](#)
- [c6gd インスタンス](#)
- [c6gn インスタンス](#)
- [c6i インスタンス](#)
- [c6id インスタンス](#)
- [c6in インスタンス](#)
- [c7a インスタンス](#)
- [c7g インスタンス](#)
- [c7gd インスタンス](#)
- [c7gn インスタンス](#)
- [c7i インスタンス](#)
- [d2 インスタンス](#)
- [d3 インスタンス](#)
- [d3en インスタンス](#)
- [g3 インスタンス](#)
- [g3s インスタンス](#)
- [g4dn インスタンス](#)
- [g5 インスタンス](#)
- [h1 インスタンス](#)
- [i2 インスタンス](#)
- [i3 インスタンス](#)
- [i3en インスタンス](#)
- [i4g インスタンス](#)

- [i4i インスタンス](#)
- [im4gn インスタンス](#)
- [is4gen インスタンス](#)
- [m1 インスタンス](#)
- [m2 インスタンス](#)
- [m3 インスタンス](#)
- [m4 インスタンス](#)
- [m5 インスタンス](#)
- [m5a インスタンス](#)
- [m5ad インスタンス](#)
- [m5d インスタンス](#)
- [m5dn インスタンス](#)
- [m5n インスタンス](#)
- [m5zn インスタンス](#)
- [m6a インスタンス](#)
- [m6g インスタンス](#)
- [m6gd インスタンス](#)
- [m6i インスタンス](#)
- [m6id インスタンス](#)
- [m6idn インスタンス](#)
- [m6in インスタンス](#)
- [m7a インスタンス](#)
- [m7g インスタンス](#)
- [m7gd インスタンス](#)
- [m7i インスタンス](#)
- [m7i-flex インスタンス](#)
- [p2 インスタンス](#)
- [p3 インスタンス](#)

- [p5 インスタンス](#)
- [r3 インスタンス](#)
- [r4 インスタンス](#)
- [r5 インスタンス](#)
- [r5a インスタンス](#)
- [r5ad インスタンス](#)
- [r5b インスタンス](#)
- [r5d インスタンス](#)
- [r5dn インスタンス](#)
- [r5n インスタンス](#)
- [r6a インスタンス](#)
- [r6g インスタンス](#)
- [r6gd インスタンス](#)
- [r6i インスタンス](#)
- [r6id インスタンス](#)
- [r6idn インスタンス](#)
- [r6in インスタンス](#)
- [r7a インスタンス](#)
- [r7g インスタンス](#)
- [r7gd インスタンス](#)
- [r7i インスタンス](#)
- [r7iz インスタンス](#)
- [x1 インスタンス](#)
- [x1e インスタンス](#)
- [x2gd インスタンス](#)
- [x2idn インスタンス](#)
- [x2iedn インスタンス](#)
- [z1d インスタンス](#)

## c1 インスタンス

### c1.medium

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	192
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	128
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

### c1.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	768
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

## c3 インスタンス

### c3.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

### c3.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

## c3.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

## c3.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

## c4 インスタンス

### c4.large

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	1152
YARN_PROXYSERVER_HEAPSIZE	1152
YARN_NODEMANAGER_HEAPSIZE	1152
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1152
HADOOP_NAMENODE_HEAPSIZE	576
HADOOP_DATANODE_HEAPSIZE	384

### c4.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

## c4.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

## c4.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064



## c4.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

## c5 インスタンス

### c5.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

## c5.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

## c5.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## c5.9xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

## c5.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

## c5.18xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

## c5.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## c5a インスタンス

### c5a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

### c5a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c5a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c5a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c5a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

## c5a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c5a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## c5ad インスタンス

## c5ad.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588



## c5ad.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c5ad.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c5ad.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c5ad.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c5ad.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c5ad.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## c5d インスタンス

### c5d.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

### c5d.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

## c5d.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## c5d.9xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

## c5d.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

## c5d.18xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

## c5d.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## c5n インスタンス

## c5n.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2304
YARN_PROXYSERVER_HEAPSIZE	2304
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2304
HADOOP_NAMENODE_HEAPSIZE	1280
HADOOP_DATANODE_HEAPSIZE	665

## c5n.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2519
YARN_PROXYSERVER_HEAPSIZE	2519
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2519
HADOOP_NAMENODE_HEAPSIZE	2355
HADOOP_DATANODE_HEAPSIZE	880

## c5n.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2949
YARN_PROXYSERVER_HEAPSIZE	2949
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2949
HADOOP_NAMENODE_HEAPSIZE	4505
HADOOP_DATANODE_HEAPSIZE	1310



## c5n.9xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

## c5n.18xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## c6a インスタンス

### c6a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

### c6a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c6a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c6a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c6a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c6a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c6a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## c6a.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## c6a.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## c6g インスタンス

### c6g.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

## c6g.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c6g.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c6g.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c6g.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324



## c6g.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c6gd インスタンス

## c6gd.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

## c6gd.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c6gd.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c6gd.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c6gd.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c6gd.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c6gn インスタンス

## c6gn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

## c6gn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c6gn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c6gn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c6gn.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c6gn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c6i インスタンス

## c6i.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

## c6i.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c6i.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075



## c6i.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c6i.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c6i.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c6i.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## c6i.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## c6id インスタンス

### c6id.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

## c6id.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c6id.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c6id.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c6id.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c6id.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c6id.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## c6id.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## c6in インスタンス

### c6in.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

## c6in.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c6in.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075



## c6in.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c6in.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c6in.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c6in.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## c6in.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## c7a インスタンス

### c7a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

## c7a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c7a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c7a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c7a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c7a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c7a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## c7a.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## c7a.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## c7g インスタンス

### c7g.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

### c7g.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762



## c7g.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c7g.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c7g.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c7g.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c7gd インスタンス

### c7gd.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

### c7gd.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c7gd.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c7gd.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c7gd.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c7gd.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c7gn インスタンス

### c7gn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

### c7gn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c7gn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c7gn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c7gn.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c7gn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949



## c7i インスタンス

### c7i.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

### c7i.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## c7i.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## c7i.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## c7i.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## c7i.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## c7i.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## c7i.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## d2 インスタンス

### d2.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### d2.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## d2.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## d2.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## d3 インスタンス

### d3.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### d3.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## d3.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## d3.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096



## d3en インスタンス

### d3en.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### d3en.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## d3en.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## d3en.6xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## d3en.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## d3en.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## g3 インスタンス

### g3.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

### g3.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## g3.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## g3s インスタンス

### g3s.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## g4dn インスタンス

### g4dn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

### g4dn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## g4dn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## g4dn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

## g4dn.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## g4dn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096



## g5 インスタンス

### g5.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### g5.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## g5.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## g5.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## g5.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## g5.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## g5.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## g5.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## h1 インスタンス

### h1.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

### h1.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## h1.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

## h1.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

## i2 インスタンス

### i2.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### i2.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## i2.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## i2.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096



## i3 インスタンス

### i3.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### i3.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## i3.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## i3.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

### i3.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

### i3en インスタンス

#### i3en.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## i3en.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## i3en.3xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

## i3en.6xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## i3en.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

## i3en.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

## i4g インスタンス

### i4g.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## i4g.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## i4g.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## i4g.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## i4g.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096



## i4i インスタンス

### i4i.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### i4i.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## i4i.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## i4i.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## i4i.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## i4i.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## i4i.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## i4i.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## im4gn インスタンス

### im4gn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### im4gn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## im4gn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## im4gn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## im4gn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## is4gen インスタンス

## is4gen.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2557
YARN_PROXYSERVER_HEAPSIZE	2557
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2557
HADOOP_NAMENODE_HEAPSIZE	2547
HADOOP_DATANODE_HEAPSIZE	919

## is4gen.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

## is4gen.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324



## is4gen.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m1 インスタンス

## m1.small

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	256
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	192
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

## m1.medium

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	384
YARN_PROXYSERVER_HEAPSIZE	192
YARN_NODEMANAGER_HEAPSIZE	256
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	256
HADOOP_NAMENODE_HEAPSIZE	384
HADOOP_DATANODE_HEAPSIZE	192

## m1.large

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	768
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

## m1.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	1024
YARN_PROXYSERVER_HEAPSIZE	512
YARN_NODEMANAGER_HEAPSIZE	768
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	2304
HADOOP_DATANODE_HEAPSIZE	384

## m2 インスタンス

## m2.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	3072
HADOOP_DATANODE_HEAPSIZE	384

## m2.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	6144
HADOOP_DATANODE_HEAPSIZE	384

## m2.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2048
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1536
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	12288
HADOOP_DATANODE_HEAPSIZE	384

## m3 インスタンス

### m3.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

### m3.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

## m4 インスタンス

### m4.large

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

### m4.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

## m4.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## m4.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## m4.10xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5365
YARN_PROXYSERVER_HEAPSIZE	5365
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5365
HADOOP_NAMENODE_HEAPSIZE	16588
HADOOP_DATANODE_HEAPSIZE	3727

## m4.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096



## m5 インスタンス

### m5.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

### m5.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## m5.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## m5.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

## m5.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## m5.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

## m5.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

## m5a インスタンス

## m5a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

## m5a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## m5a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## m5a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

## m5a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## m5a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

## m5a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

## m5ad インスタンス

### m5ad.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m5ad.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075



## m5ad.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m5ad.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m5ad.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m5ad.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m5ad.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m5d インスタンス

## m5d.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

## m5d.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## m5d.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## m5d.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

## m5d.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## m5d.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

## m5d.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

## m5dn インスタンス

### m5dn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m5dn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m5dn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m5dn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949



## m5dn.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m5dn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m5dn.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m5n インスタンス

## m5n.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## m5n.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m5n.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m5n.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m5n.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m5n.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m5n.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m5zn インスタンス

### m5zn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

### m5zn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m5zn.3xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

## m5zn.6xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

## m5zn.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m6a インスタンス

## m6a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762



## m6a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m6a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m6a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m6a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m6a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m6a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m6a.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## m6a.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## m6g インスタンス

### m6g.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m6g.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m6g.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m6g.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m6g.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

## m6g.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m6gd インスタンス

### m6gd.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m6gd.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075



## m6gd.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m6gd.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m6gd.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

## m6gd.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m6i インスタンス

### m6i.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m6i.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m6i.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m6i.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m6i.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

## m6i.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m6i.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m6i.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## m6id インスタンス

### m6id.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m6id.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m6id.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m6id.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949



## m6id.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m6id.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m6id.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m6id.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## m6idn インスタンス

### m6idn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m6idn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m6idn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m6idn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m6idn.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m6idn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m6idn.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m6idn.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## m6in インスタンス

### m6in.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m6in.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m6in.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m6in.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949



## m6in.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

## m6in.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m6in.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m6in.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## m7a インスタンス

### m7a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

### m7a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m7a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m7a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m7a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m7a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m7a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## m7a.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## m7a.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## m7g インスタンス

## m7g.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## m7g.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m7g.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699



## m7g.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m7g.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m7g.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m7gd インスタンス

## m7gd.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## m7gd.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m7gd.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m7gd.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m7gd.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

## m7gd.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m7i インスタンス

## m7i.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## m7i.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m7i.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m7i.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## m7i.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

## m7i.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## m7i.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096



## m7i.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## m7i-flex インスタンス

## m7i-flex.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

## m7i-flex.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## m7i-flex.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## m7i-flex.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## p2 インスタンス

## p2.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## p2.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## p2.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## p3 インスタンス

### p3.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

### p3.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## p3.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## p5 インスタンス

## p5.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

## r3 インスタンス

### r3.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r3.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r3.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r3.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096



## r4 インスタンス

### r4.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r4.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r4.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r4.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r4.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r5 インスタンス

## r5.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## r5.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## r5.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

## r5.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

## r5.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

## r5.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

## r5.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

## r5a インスタンス

### r5a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

### r5a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## r5a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

## r5a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096



## r5a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

## r5a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

## r5a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

## r5ad インスタンス

## r5ad.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## r5ad.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r5ad.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r5ad.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r5ad.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r5ad.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12247
YARN_PROXYSERVER_HEAPSIZE	12247
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12247
HADOOP_NAMENODE_HEAPSIZE	50995
HADOOP_DATANODE_HEAPSIZE	4096

## r5ad.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r5b インスタンス

### r5b.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r5b.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r5b.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r5b.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r5b.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r5b.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096



## r5b.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r5d インスタンス

## r5d.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## r5d.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## r5d.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

## r5d.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

## r5d.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

## r5d.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

## r5d.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

## r5dn インスタンス

### r5dn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r5dn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r5dn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r5dn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r5dn.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r5dn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r5dn.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r5n インスタンス

## r5n.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075



## r5n.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r5n.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r5n.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r5n.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r5n.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r5n.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r6a インスタンス

### r6a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r6a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r6a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r6a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r6a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r6a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r6a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r6a.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## r6a.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

## r6g インスタンス

## r6g.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075



## r6g.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r6g.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r6g.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r6g.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r6g.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r6gd インスタンス

## r6gd.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## r6gd.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r6gd.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r6gd.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r6gd.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r6gd.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r6i インスタンス

## r6i.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## r6i.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r6i.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r6i.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r6i.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096



## r6i.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r6i.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r6i.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

## r6id インスタンス

## r6id.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## r6id.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r6id.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r6id.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r6id.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r6id.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r6id.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r6id.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## r6idn インスタンス

## r6idn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## r6idn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r6idn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r6idn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r6idn.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096



## r6idn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r6idn.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r6idn.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## r6in インスタンス

## r6in.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## r6in.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r6in.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r6in.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r6in.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r6in.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r6in.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r6in.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## r7a インスタンス

## r7a.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

## r7a.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r7a.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r7a.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r7a.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096



## r7a.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r7a.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r7a.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## r7a.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

## r7g インスタンス

### r7g.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r7g.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r7g.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r7g.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r7g.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r7g.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r7gd インスタンス

### r7gd.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r7gd.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r7gd.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r7gd.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r7gd.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r7gd.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096



## r7i インスタンス

### r7i.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r7i.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r7i.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r7i.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r7i.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r7i.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r7i.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## r7i.48xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

## r7iz インスタンス

### r7iz.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

### r7iz.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## r7iz.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## r7iz.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## r7iz.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

## r7iz.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## r7iz.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

## x1 インスタンス

## x1.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096



## x1.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

## x1e インスタンス

## x1e.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4520
YARN_PROXYSERVER_HEAPSIZE	4520
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4520
HADOOP_NAMENODE_HEAPSIZE	12364
HADOOP_DATANODE_HEAPSIZE	2882

## x1e.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6952
YARN_PROXYSERVER_HEAPSIZE	6952
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6952
HADOOP_NAMENODE_HEAPSIZE	24524
HADOOP_DATANODE_HEAPSIZE	4096

## x1e.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	11816
YARN_PROXYSERVER_HEAPSIZE	11816
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	11816
HADOOP_NAMENODE_HEAPSIZE	48844
HADOOP_DATANODE_HEAPSIZE	4096

## x1e.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

## x1e.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

## x1e.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	79912
YARN_PROXYSERVER_HEAPSIZE	79912
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	79912
HADOOP_NAMENODE_HEAPSIZE	389324
HADOOP_DATANODE_HEAPSIZE	4096

## x2gd インスタンス

## x2gd.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

## x2gd.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

## x2gd.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## x2gd.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## x2gd.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

## x2gd.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## x2idn インスタンス

## x2idn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## x2idn.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

## x2idn.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096



## x2iedn インスタンス

### x2iedn.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

### x2iedn.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

## x2iedn.4xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

## x2iedn.8xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

## x2iedn.16xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

## x2iedn.24xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

## x2iedn.32xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	82042
YARN_PROXYSERVER_HEAPSIZE	82042
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	82042
HADOOP_NAMENODE_HEAPSIZE	399974
HADOOP_DATANODE_HEAPSIZE	4096

## z1d インスタンス

## z1d.xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

## z1d.2xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

## z1d.3xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

## z1d.6xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

## z1d.12xlarge

パラメータ	Value
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

## HDFS 構成

次の表に、Hadoop Distributed File System ( HDFS ) のデフォルトのパラメータとその設定を示します。hdfs-site 設定分類を使用して、これらの値を変更できます。詳細については、「[アプリケーションの設定](#)」を参照してください。

### ⚠ Warning

1. ノードが 4 つ未満のクラスターで dfs.replication を 1 に設定すると、単一ノードがダウンした場合に HDFS データが失われる可能性があります。クラスターに HDFS ストレージがある場合は、データ損失を避けるため、本番稼働ワークロード用に少なくとも 4 つのコアノードでクラスターを構成することをお勧めします。
2. Amazon EMR では、クラスターはコアノードを dfs.replication 未満にスケールすることはできません。例えば、dfs.replication = 2 の場合、コアノードの最小数は 2 です。
3. マネージドスケールリングや自動スケールリングを使用する場合や、クラスターのサイズを手動で変更する場合は、dfs.replication を 2 以上に設定することをお勧めします。

パラメータ	定義	デフォルト値
dfs.block.size	HDFS ブロックのサイズ。HDFS に格納されているデータを操作するとき、分割サイズは原則として HDFS ブロックのサイズです。値を大きくするとタスクの粒度は小さくなりますが、クラスターへの負荷も低下します。NameNode	134217728 ( 128 MB )
dfs.replication	耐久性のために保存する各ブロックのコピー数。Amazon EMR は、クラスターがプロビジョニングされるコアノードの数に基づいてこの値を設定します。必要に応じて値を調整してください。デフォルト値を上書きするには、hdfs-site 分類を使用します。	1 (コアノードが 4 個未満でプロビジョニングされているクラスターの場合)  2 (コアノードが 10 個未満でプロビジョニングされ

パラメータ	定義	デフォルト値
		ているクラスターの場合)  3 (それ以外のすべてのクラスターの場合)

## Amazon EMR における HDFS での透過的暗号化

透過性暗号化は、お客様が定義する HDFS パスである HDFS 暗号化ゾーンを通じて実装されます。各暗号化ゾーンには独自のキーがあり、`hdfs-site` 設定分類を使用して指定されたキーサーバーに保存されます。

Amazon EMR リリースバージョン 4.8.0 から、Amazon EMR セキュリティ設定を使用して、クラスターのデータ暗号化設定をより簡単に行うことができます。セキュリティ設定は、Amazon Elastic Block Store (Amazon EBS) ストレージボリュームおよび Amazon S3 の EMRFS データで伝送時のデータと保管時のデータに対するセキュリティを有効にするオプションを提供します。詳細については、「Amazon EMR 管理ガイド」の「[データを伝送中と保管時に暗号化](#)」を参照してください。

Amazon EMR はデフォルトで Hadoop KMS を使用しますが、KeyProvider API オペレーションを実装する別の KMS を使用できます。HDFS 暗号化ゾーンの各ファイルには、独自のデータ暗号化キーがあり、これは暗号化ゾーンキーで暗号化されます。暗号化および復号アクティビティはクライアントでのみ発生するため、HDFS データは暗号化ゾーンに書き込まれると end-to-end (保管中および転送中の) 暗号化されます。

暗号化ゾーン間でファイルを移動したり、暗号化ゾーンから暗号化されていないパスにファイルを移動することはできません。

NameNode および HDFS クライアントは、KeyProvider API オペレーションを通じて Hadoop KMS (または設定した代替 KMS) とやり取りします。KMS はバックアップキーストアで暗号化キーを保存します。また、Amazon EMR には JCE 無制限強度のポリシーが含まれているため、必要な長さでキーを作成できます。

詳細については、Hadoop のドキュメントの「[Transparent encryption in HDFS \(HDFS での透過的暗号化\)](#)」を参照してください。



**Note**

Amazon EMR では、Hadoop KMS における HTTPS 経由の KMS はデフォルトで有効になりません。HTTPS 経由の KMS を有効にする方法については、[Hadoop KMS のドキュメント](#)を参照してください

## HDFS 透過的暗号化を設定する

キーを作成し、暗号化ゾーンを追加することで、Amazon EMR で透過的暗号化を設定できます。これはいくつかの方法で行うことができます。

- クラスターの作成時に Amazon EMR 設定 API オペレーションを使用する
- `command-runner.jar` で Hadoop JAR のステップを使用する
- Hadoop クラスターのマスターノードにログインして、`hadoop key` と `hdfs crypto` コマンドラインクライアントを使用する
- Hadoop KMS と HDFS で REST API を使用する

REST API の詳細については、Hadoop KMS と HDFS の各ドキュメントを参照してください。

CLI を使用してクラスター作成時に暗号化ゾーンとそのキーを作成するには

設定 API オペレーションの `hdfs-encryption-zones` 分類により、クラスターの作成時にキー名と暗号化ゾーンを指定できます。Amazon EMR はクラスターの Hadoop KMS でこのキーを作成し、暗号化ゾーンを設定します。

- 次のコマンドを使用してクラスターを作成します。

```
aws emr create-cluster --release-label emr-7.1.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/
mybucket/myfolder/myConfig.json
```

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

myConfig.json:

```
[
  {
    "Classification": "hdfs-encryption-zones",
    "Properties": {
      "/myHDFSPath1": "path1_key",
      "/myHDFSPath2": "path2_key"
    }
  }
]
```

マスターノードで手動で暗号化ゾーンおよびそのキーを作成するには

1. Amazon EMR リリース 4.1.0 以降を使用してクラスターを起動します。
2. SSH を使用してクラスターのマスターノードに接続します。
3. Hadoop KMS 内でキーを作成します。

```
$ hadoop key create path2_key
path2_key has been successfully created with options Options{cipher='AES/CTR/
NoPadding', bitLength=256, description='null', attributes=null}.
KMSClientProvider[http://ip-x-x-x-x.ec2.internal:16000/kms/v1/] has been updated.
```

**Important**

Hadoop KMS では、キー名が小文字である必要があります。大文字が含まれているキーを使用すると、起動中にクラスターが失敗します。

4. HDFS で暗号化ゾーンのパスを作成します。

```
$ hadoop fs -mkdir /myHDFSPath2
```

## 5. 作成したキーを使用して HDFS パスを暗号化ゾーンにします。

```
$ hdfs crypto -createZone -keyName path2_key -path /myHDFSPath2
Added encryption zone /myHDFSPath2
```

を使用して暗号化ゾーンとそのキーを手動で作成するには AWS CLI

- 次のコマンドで KMS キーおよび暗号化ゾーンを作成するステップを追加します。

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Create
  First Hadoop KMS Key",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[/
bin/bash,-c,"\"hadoop key create path1_key\""] \
Type=CUSTOM_JAR,Name="Create First Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create First Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
-keyName path1_key -path /myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop KMS Key",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop key create
path2_key\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath2\""] \
Type=CUSTOM_JAR,Name="Create Second Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
-keyName path2_key -path /myHDFSPath2\""]
```

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

## HDFS 透過的暗号化の考慮事項

ベストプラクティスとして、アプリケーションがファイルを書き込む可能性がある各アプリケーション用に暗号化ゾーンを作成します。また、設定 API の `hdfs-encryption-zones` 分類を使用してすべての HDFS を暗号化し、ルートパス (/) を暗号化ゾーンとして指定することもできます。

### Hadoop キー管理サーバー

[Hadoop KMS](#) は、Hadoop クラスター用の暗号化サービスを実装する機能を提供するキー管理サーバーであり、[Amazon EMR における HDFS での透過的暗号化](#) のキーベンダーとして機能することができます。Amazon EMR の Hadoop KMS は、EMR クラスターの起動時に Hadoop アプリケーションを選択するときにデフォルトでインストールされ、有効になります。Hadoop KMS は、一時キャッシュの場合を除き、キー自体を保存しません。Hadoop KMS は、キープロバイダーとクライアントトラスティの間のバックアップストアへのプロキシとして機能します。これはキーストアではありません。Hadoop KMS 用に作成されるデフォルトのキーストアは、Java Cryptography Extension KeyStore (JCEKS) です。JCE の無制限強度のポリシーも含まれるため、必要な長さのキーを作成できます。Hadoop KMS は、HDFS など他のクライアントアプリケーションとは関係なく、キーとキーオペレーションへのアクセスを制御する一連の ACL もサポートします。Amazon EMR のデフォルトのキー長は 256 ビットです。

Hadoop KMS を設定するには、`hadoop-kms-site` 分類を使用して設定を変更します。ACL を設定するには、分類の `kms-acls` を使用します。

詳細については、[Hadoop KMS のドキュメント](#) を参照してください。Hadoop KMS は、Hadoop HDFS 透過的暗号化で使用されます。HDFS 透過的暗号化の詳細については、Apache Hadoop のドキュメントの「[HDFS 透過的暗号化](#)」トピックを参照してください。

#### Note

Amazon EMR では、Hadoop KMS における HTTPS 経由の KMS はデフォルトで有効になりません。HTTPS 経由で KMS を有効にする方法については、[Hadoop KMS のドキュメント](#) を参照してください。

#### Important

Hadoop KMS では、キー名が小文字である必要があります。大文字が含まれているキーを使用すると、起動中にクラスターが失敗します。

## Amazon EMR での Hadoop KMS の設定

Amazon EMR リリースバージョン 4.6.0 以降を使用すると、kms-http-port は 9700、kms-admin-port は 9701 になります。

Amazon EMR リリースの設定 API を使用して、クラスター作成時に Hadoop KMS を設定できます。次の設定オブジェクト分類が Hadoop KMS で使用できます。

### Hadoop KMS 設定分類

分類	Filename
hadoop-kms-site	kms-site.xml
hadoop-kms-acls	kms-acls.xml
hadoop-kms-env	kms-env.sh
hadoop-kms-log4j	kms-log4j.properties

CLI を使用して Hadoop KMS の ACL を設定するには

- Hadoop KMS と ACL でクラスターを作成するには、次のコマンドを使用します。

```
aws emr create-cluster --release-label emr-7.1.0 --instance-type m5.xlarge --instance-count 2 \  
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/  
mybucket/myfolder/myConfig.json
```

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

myConfig.json:

```
[  
  {
```

```
"Classification": "hadoop-kms-acls",
"Properties": {
  "hadoop.kms.blacklist.CREATE": "hdfs,foo,myBannedUser",
  "hadoop.kms.acl.ROLLOVER": "myAllowedUser"
}
}
```

CLI を使用して Hadoop KMS のキャッシュを無効にするには

- 次のコマンドを使用し、Hadoop KMS `hadoop.kms.cache.enable` を `false` に設定してクラスターを作成します。

```
aws emr create-cluster --release-label emr-7.1.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/
mybucket/myfolder/myConfig.json
```

#### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

myConfig.json:

```
[
  {
    "Classification": "hadoop-kms-site",
    "Properties": {
      "hadoop.kms.cache.enable": "false"
    }
  }
]
```

CLI を使用して `kms-env.sh` スクリプトで環境変数を設定するには

- `kms-env.sh` 設定を通じて `hadoop-kms-env` で設定を変更できます。次のコマンドを使用し、Hadoop KMS でクラスターを作成します。

```
aws emr create-cluster --release-label emr-7.1.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/
mybucket/myfolder/myConfig.json
```

#### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

myConfig.json:

```
[
  {
    "Classification": "hadoop-kms-env",
    "Properties": {
    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "JAVA_LIBRARY_PATH": "/path/to/files",
          "KMS_SSL_KEYSTORE_FILE": "/non/Default/Path/.keystore",
          "KMS_SSL_KEYSTORE_PASS": "myPass"
        },
        "Configurations": [
        ]
      }
    ]
  }
]
```

Hadoop KMS の設定については、[Hadoop KMS のドキュメント](#)を参照してください。

## 複数のマスターノードを持つ EMR クラスターでの HDFS 透過的暗号化

[Apache Ranger](#) KMS は、HDFS での透過的暗号化のために複数のプライマリノードを持つ Amazon EMR クラスターで使用されます。

Apache Ranger KMS は、ルートキーと暗号化ゾーン (EZ) キーを複数のプライマリノードを持つ Amazon EMR クラスターの Amazon RDS に保存します。複数のプライマリノードを持つ Amazon EMR クラスターで HDFS の透過的暗号化を有効にするには、次の設定を指定する必要があります。

- Ranger KMS のルートキーおよび EZ キーを保存するための Amazon RDS または独自の MySQL サーバー接続 URL
- MySQL のユーザー名とパスワード
- Ranger KMS ルートキーのパスワード
- MySQL サーバーへの SSL 接続用の認証機関 (CA) PEM ファイル

次の例に示すように、`ranger-kms-dbks-site` 分類と `ranger-kms-db-ca` 分類を使用して、これらの設定を提供できます。

```
[
  {
    "Classification": "ranger-kms-dbks-site",
    "Properties": {
      "ranger.ks.jpa.jdbc.url": "jdbc:log4jdbc:mysql://mysql-host-url.xx-xxx-1.xxx.amazonaws.com:3306/rangerkms",
      "ranger.ks.jpa.jdbc.user": "mysql-user-name",
      "ranger.ks.jpa.jdbc.password": "mysql-password",
      "ranger.db.encrypt.key.password": "password-for-encrypting-a-master-key"
    }
  },
  {
    "Classification": "ranger-kms-db-ca",
    "Properties": {
      "ranger.kms.trust.ca.file.s3.url": "s3://rds-downloads/rds-ca-2019-root.pem"
    }
  }
]
```

以下は、Apache Ranger KMS の設定オブジェクト分類です。



## Hadoop KMS 設定分類

分類	説明
ranger-kms-dbks-site	Ranger KMS の dbks-site.xml ファイルの値を変更します。
ranger-kms-site	Ranger KMS の ranger-kms-site.xml ファイルの値を変更します。
ranger-kms-env	Ranger KMS 環境の値を変更します。
ranger-kms-log4j	Ranger KMS の kms-log4j.properties ファイルの値を変更します。
ranger-kms-db-ca	Ranger KMS を使用した MySQL SSL 接続用の S3 の CA ファイルの値を変更します。

### 考慮事項

- セキュリティ向上のため、Amazon RDS インスタンスを暗号化することを強くお勧めします。詳細については、「[Amazon RDS リソースの暗号化の概要](#)」を参照してください。
- セキュリティのハードルを上げるため、複数のプライマリノードを持つ各 Amazon EMR クラスターに別々の MySQL データベースを使用することを強くお勧めします。
- 複数のプライマリノードを持つ Amazon EMR クラスターの HDFS において透過的な暗号化を設定するには、クラスターの作成時に `hdfs-encryption-zones` 分類を指定する必要があります。指定しない場合、Ranger KMS は設定も起動もしません。複数のプライマリノードを持つ Amazon EMR クラスターでは、`hdfs-encryption-zones` 分類の再設定、または実行中のクラスターでの Hadoop KMS 設定分類はサポートされていません。

## Hadoop アプリケーションの作成または実行

### トピック

- [Amazon EMR を使用してバイナリを構築する](#)
- [ストリーミングでのデータ処理](#)
- [カスタム JAR でのデータの処理](#)

## Amazon EMR を使用してバイナリを構築する

クラスターで使用するプログラムをコンパイルするためのビルド環境として Amazon EMR を使用できます。Amazon EMR で使用するプログラムのコンパイルは、Amazon EMR が使用しているバージョンと同じ Linux が実行されているシステムで行わなければなりません。32 ビットバージョンについては、32 ビットマシンまたは 32 ビットクロスコンパイルオプションをオンにして、コンパイルしておく必要があります。64 ビットバージョンについては、64 ビットマシンまたは 64 ビットクロスコンパイルオプションをオンにして、コンパイルしておく必要があります。EC2 インスタンスのバージョンの詳細については、「Amazon EMR 管理ガイド」の「[EC2 インスタンスを計画し構成する](#)」を参照してください。C++、Python、C# などのプログラミング言語がサポートされています。

次の表は、Amazon EMR を使用してアプリケーションを構築し、テストするステップを簡単に示しています。

### モジュール構築のプロセス

- 1 クラスターのマスターノードに接続します。
- 2 ソースファイルをマスターノードにコピーします。
- 3 バイナリを構築し、必要に応じて最適化します。
- 4 バイナリをマスターノードから Amazon S3 にコピーします。

各ステップの詳細については、以降のセクションで説明します。

### クラスターのマスターノードに接続するには

- 「Amazon EMR 管理ガイド」の「[SSH を使用してマスターノードに接続する](#)」にある手順を実行してください。

### ソースファイルをマスターノードにコピーするには

1. Amazon S3 バケットにソース ファイルを置きます。バケットを作成する方法とデータを Amazon S3 に移動する方法については、「[Amazon Simple Storage Service ユーザーガイド](#)」を参照してください。
2. 以下のようなコマンドを入力して、ソースファイルを入れるフォルダーを Hadoop クラスター上に作成します:

```
mkdir SourceFiles
```

3. 次のようなコマンドを入力して、ソースファイルを Amazon S3 からマスターノードにコピーします。

```
hadoop fs -get s3://mybucket/SourceFiles SourceFiles
```

## バイナリを構築し、必要に応じて最適化する

バイナリを構築する方法はさまざまな要因によって決まります。ご利用の構築ツールの手順に従って、環境をセットアップして設定してください。Hadoop システム仕様コマンドを使用すると、構築環境をインストールする方法を判断する際に必要なクラスター情報を入手できます。

### システム仕様を確認するには

- バイナリを構築する際に使用するアーキテクチャを確認するには、以下のコマンドを使用します。
  - a. Debian のバージョンを確認するには、以下のコマンドを入力します:

```
master$ cat /etc/issue
```

出力は次の例のようになります。

```
Debian GNU/Linux 5.0
```

- b. パブリック DNS 名とプロセッサのサイズを確認するには、以下のコマンドを入力します。

```
master$ uname -a
```

出力は次の例のようになります。

```
Linux domU-12-31-39-17-29-39.compute-1.internal 2.6.21.7-2.fc8xen #1 SMP Fri  
Feb 15 12:34:28 EST 2008 x86_64 GNU/Linux
```

- c. プロセッサ速度を確認するには、以下のコマンドを入力します:

```
master$ cat /proc/cpuinfo
```

出力は次の例のようになります。

```
processor : 0
vendor_id : GenuineIntel
model name : Intel(R) Xeon(R) CPU E5430 @ 2.66GHz
flags : fpu tsc msr pae mce cx8 apic mca cmov pat pse36 clflush dts acpi mmx
      fxsr sse sse2 ss ht tm syscall nx lm constant_tsc pni monitor ds_cpl vmx est
      tm2 ssse3 cx16 xtpr cda lahf_lm
...
```

バイナリを構築したら、ファイルを Amazon S3 にコピーできます。

バイナリをマスターノードから Amazon S3 にコピーするには

- 次のコマンドを入力して、バイナリを Amazon S3 バケットにコピーします。

```
hadoop fs -put BinaryFiles s3://mybucket/BinaryDestination
```

## ストリーミングでのデータ処理

Hadoop ストリーミングは、Java 以外の言語で実行可能ファイルを開発 MapReduce できる Hadoop に付属するユーティリティです。JAR ファイルの形式でストリーミングが実装されるため、Amazon EMR API から、またはコマンドラインから標準の JAR ファイルの用に実行できます。

このセクションでは、Amazon EMR でストリーミングを使用する方法について説明します。

### Note

Apache Hadoop Streaming は独立したツールです。そのため、ここでは、その関数とパラメータすべてを説明するわけではありません。Hadoop ストリーミングの詳細については、<http://hadoop.apache.org/docs/stable/hadoop-streaming/HadoopStreaming.html> を参照してください。

## Hadoop ストリーミングユーティリティの使用

このセクションでは、Hadoop のストリーミングユーティリティの使用方法について説明します。

### Hadoop プロセス

1	自由にプログラム言語を選択して、実行可能なマッパーとリデューサーを作成します。  Hadoop のドキュメントの指示に従ってストリーミング実行ファイルを書いてください。プログラムは、標準出力による標準入出力データからの入力を読み取れるようにする必要があります。デフォルトでは、入出力の各行は 1 つのレコードに対応し、各行の最初のタブはキーと値を区切ります。
2	ローカルで実行ファイルをテストしてから、Amazon S3 へアップロードします。
3	Amazon EMR コマンドラインインターフェイスまたは Amazon EMR コンソールを使用してアプリケーションを実行します。

各マッパースクリプトは、クラスター内の独立したプロセスとして実行されます。各リデューサー実行ファイルは、マッパー実行ファイルの出力をジョブフローによるデータ出力にします。

input、output、mapper、および reducer パラメータは、ほとんどのストリーミングアプリケーションで必要とされます。以下の表は、これらのパラメータと他のオプションパラメータについて説明します。

パラメータ	説明	必須
-input	Amazon S3 上の入力データの場所。  型: 文字列  デフォルト: なし  制約: URI。プロトコルの指定がなければ、クラスターのデフォルトファイルシステムが使用されます。	はい
-output	Amazon EMR が処理されたデータをアップロードする Amazon S3 上の場所。	はい

パラメータ	説明	必須
	<p>型: 文字列</p> <p>デフォルト: なし</p> <p>制約: URI</p> <p>デフォルト: 場所の指定がなければ、Amazon EMR は input で指定された場所にデータをアップロードします。</p>	
-mapper	<p>マッパー実行ファイルの名前。</p> <p>型: 文字列</p> <p>デフォルト: なし</p>	はい
-reducer	<p>リデューサー実行ファイルの名前。</p> <p>型: 文字列</p> <p>デフォルト: なし</p>	はい
-cacheFile	<p>( 主としてパフォーマンス向上のために ) Hadoop がローカル作業ディレクトリにコピーするファイルが格納されている Amazon S3 の場所。</p> <p>型: 文字列</p> <p>デフォルト: なし</p> <p>制約: [URI]#[作業ディレクトリに作成するシンボリックリンク名]</p>	いいえ

パラメータ	説明	必須
-cacheArchive	<p>作業ディレクトリに抽出する JAR ファイル</p> <p>型: 文字列</p> <p>デフォルト: なし</p> <p>制約: [URI]#[作業ディレクトリに作成するシンボリックリンクディレクトリ名]</p>	いいえ
-combiner	<p>結果の結合</p> <p>型: 文字列</p> <p>デフォルト: なし</p> <p>制約: Java クラス名</p>	いいえ

次のコード例は、Python で作成されたマッパーの実行可能ファイルです。このスクリプトは WordCount サンプルアプリケーションの一部です。

```
#!/usr/bin/python
import sys

def main(argv):
    line = sys.stdin.readline()
    try:
        while line:
            line = line.rstrip()
            words = line.split()
            for word in words:
                print "LongValueSum:" + word + "\t" + "1"
            line = sys.stdin.readline()
    except "end of file":
        return None
if __name__ == "__main__":
    main(sys.argv)
```

## ストリーミングステップの送信

このセクションでは、クラスターにストリーミングステップを送信するための基本について説明します。ストリーミングアプリケーションは、標準入力から入力を読み取って、スクリプトまたは実行可能ファイル (マッパーと呼ばれます) を各入力に対して実行します。各入力の結果は、ローカル (通常は Hadoop Distributed File System (HDFS) パーティション) に保存されます。すべての入力がマッパーによって処理されたら、そのマッパーの結果は、2 目目のスクリプトまたは実行可能ファイル (リデューサーと呼ばれます) によって処理されます。リデューサーの結果は標準出力に送信されます。一連のストリーミングステップをつなげて、あるステップの出力を別のステップの入力にすることができます。

マッパーとリデューサーはそれぞれファイルとして参照できます。また、Java クラスを指定することもできます。マッパーおよびリデューサーは、Ruby、Perl、Python、PHP、Bash など、サポートされている任意の言語で実装できます。

### コンソールを使用したストリーミングステップの送信

この例では、Amazon EMR コンソールを使用し、実行中のクラスターにストリーミングステップを送信する方法について説明します。

ストリーミングステップを送信するには

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Cluster List (クラスターリスト)] で、クラスターの名前を選択します。
3. [Steps (ステップ)] セクションまでスクロールして展開し、[Add step (ステップの追加)] を選択します。
4. [Add Step] ダイアログボックスで、次のようにします。
  - [Step type] は [Streaming program] を選択します。
  - [Name] では、デフォルト名 ( Streaming program ) を使用するか、または新しい名前を入力します。
  - [Mapper (マッパー)] では、Hadoop のマッパークラスの場所、または Python プログラムなどのマッパー実行ファイルが存在する S3 バケットを入力するか参照します。パス値は *BucketName/path* の形式である必要があります *MapperExecutable*。
  - [Reducer (リデューサー)] では、Hadoop のリデューサークラスの場所、または Python プログラムなどのリデューサー実行ファイルが存在する S3 バケットを入力するか参照します。パス値は *BucketName/path* の形式である必要があります *MapperExecutable*。Amazon EMR



は、特別な aggregate キーワードをサポートしています。詳細については、Hadoop が提供する Aggregate ライブラリを参照してください。

- [Input S3 location] では、入力データの場所を入力または参照します。
  - [Output S3 location] (S3 の場所の出力) で、Amazon S3 出力バケットの名前を入力するか、参照します。
  - [Arguments (引数)] は、フィールドを空のままにします。
  - [Action on failure (失敗時の操作)] では、デフォルトのオプション ([Continue (続行)]) を使用します。
5. [追加] を選択します。ステップは、[Pending] というステータスでコンソールに表示されます。
  6. ステップが実行されると、ステータスは [Pending (保留中)] から [Running (実行中)]、[Running (実行中)] から [完了済み] に変更されます。ステータスを更新するには、[Actions] 列の上にある [Refresh] アイコンを選択します。

## AWS CLI

これらの例は、を使用してクラスター AWS CLI を作成し、ストリーミングステップを送信する方法を示しています。

を使用してクラスターを作成し、ストリーミングステップを送信するには AWS CLI

- を使用してクラスターを作成し、ストリーミングステップを送信するには AWS CLI、次のコマンドを入力し、*myKey* を EC2 キーペアの名前に置き換えます。--files の引数はスクリプトの場所の Amazon S3 パスで、-mapper および -reducer の引数はそれぞれのスクリプトファイルの名前にする必要があります。

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.1.0 --
applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=STREAMING,Name="Streaming Program",ActionOnFailure=CONTINUE,Args=[--
files,pathtoscripts,-mapper,mapperscript,-reducer,reducerscript,aggregate,-
input,pathtoinputdata,-output,pathtooutputbucket]
```

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

--instance-groups パラメータを使用せずにインスタンス数を指定すると、1つのマスターノードが起動され、残りのインスタンスはコアノードとして起動されます。すべてのノードで、コマンドで指定したインスタンスタイプが使用されます。

**Note**

以前にデフォルトの Amazon EMR サービスロールと EC2 インスタンスプロファイルを作成していない場合は、「aws emr create-default-roles」と入力してそれらを作成してから、create-cluster サブコマンドを入力します。

での Amazon EMR コマンドの使用の詳細については AWS CLI、「」を参照してください <https://docs.aws.amazon.com/cli/latest/reference/emr>。

## カスタム JAR でのデータの処理

カスタム JAR では、Amazon S3 にアップロードできるコンパイル済みの Java プログラムを実行します。起動する Hadoop のバージョンに合わせてプログラムをコンパイルし、Amazon EMR クラスターに CUSTOM\_JAR ステップを送信する必要があります。JAR ファイルのコンパイル方法の詳細については、「[Amazon EMR を使用してバイナリを構築する](#)」を参照してください。

Hadoop MapReduce アプリケーションの構築の詳細については、Apache Hadoop ドキュメントの [MapReduce チュートリアル](#) を参照してください。

### トピック

- [カスタム JAR ステップの送信](#)

## カスタム JAR ステップの送信

カスタム JAR では、Amazon S3 にアップロードできるコンパイル済みの Java プログラムを実行します。起動する Hadoop のバージョンに合わせてプログラムをコンパイルし、Amazon EMR クラスターに CUSTOM\_JAR ステップを送信する必要があります。JAR ファイルのコンパイル方法の詳細については、「[Amazon EMR を使用してバイナリを構築する](#)」を参照してください。

Hadoop MapReduce アプリケーションの構築の詳細については、Apache Hadoop ドキュメントの[MapReduce チュートリアル](#)を参照してください。

このセクションでは、Amazon EMR でのカスタム JAR ステップの送信の基本情報について説明します。カスタム JAR ステップを送信する場合は、データを処理するスクリプトを Java プログラミング言語で記述できます。

### コンソールを使用したカスタム JAR ステップの送信

この例では、Amazon EMR コンソールを使用して、実行中のクラスターにカスタム JAR ステップを送信する方法について説明します。

コンソールを使用してカスタム JAR ステップを送信するには

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Cluster List (クラスターリスト)] で、クラスターの名前を選択します。
3. [Steps (ステップ)] セクションまでスクロールして展開し、[Add step (ステップの追加)] を選択します。
4. [Add Step (ステップの追加)] ダイアログで:
  - [Step type] で、[Custom JAR] を選択します。
  - [Name] では、デフォルト名 ( Custom JAR ) を使用するか、または新しい名前を入力します。
  - [JAR S3 location] では、JAR ファイルの場所を入力または参照します。JAR の場所は、S3 へのパスまたはクラスパス内の完全修飾 Java クラスです。
  - [Arguments (引数)] では、必要な引数をスペース区切りの文字列として入力するか、フィールドを空のままにします。
  - [Action on failure (失敗時の操作)] では、デフォルトのオプション ([Continue (続行)]) を使用します。
5. [追加] を選択します。ステップは、[Pending] というステータスでコンソールに表示されます。

6. ステップが実行されると、ステータスは [Pending (保留中)] から [Running (実行中)]、[Running (実行中)] から [完了済み] に変更されます。ステータスを更新するには、[Actions] 列の上にある [Refresh] アイコンを選択します。

を使用してクラスターを起動し、カスタム JAR ステップを送信する AWS CLI

を使用してクラスターを起動し、カスタム JAR ステップを送信するには AWS CLI

クラスターを起動し、を使用してカスタム JAR ステップを送信するには AWS CLI、`--steps` パラメータを指定して `create-cluster` サブコマンドを入力します。

- クラスターを起動し、カスタム JAR ステップを送信するには、次のコマンドを入力して *myKey* を EC2 キーペア名に置き換え、*mybucket* をバケット名に置き換えます。

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.1.0 \  
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \  
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \  
--steps Type=CUSTOM_JAR,Name="Custom JAR  
Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbucket"]
```

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

`--instance-groups` パラメータを使用せずにインスタンス数を指定すると、1 つのプライマリノードが起動され、残りのインスタンスはコアノードとして起動されます。すべてのノードで、コマンドで指定したインスタンスタイプが使用されます。

#### Note

以前にデフォルトの Amazon EMR サービスロールと EC2 インスタンスプロファイルを作成していない場合は、「aws emr create-default-roles」と入力してそれらを作成してから、`create-cluster` サブコマンドを入力します。

での Amazon EMR コマンドの使用の詳細については AWS CLI、「」を参照してください <https://docs.aws.amazon.com/cli/latest/reference/emr>。

## サードパーティーへの依存関係

プログラムで使用するクラス MapReduce パス JARsに を含める必要がある場合があります。これを行うための 2 つのオプションがあります。

- 「`--libjars s3://URI_to_JAR`」の手順で、ステップオプションに [を使用してクラスターを起動し、カスタム JAR ステップを送信する AWS CLI](#) を含めます。
- `mapred-site.xml` で `mapreduce.application.classpath` 設定を変更して、クラスターを起動します。 `mapred-site` 設定分類を使用します。を使用して ステップでクラスターを作成するには AWS CLI、次のようになります。

```
aws emr create-cluster --release-label emr-7.1.0 \  
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \  
--instance-type m5.xlarge --instance-count 2 --ec2-attributes KeyName=myKey \  
--steps Type=CUSTOM_JAR,Name="Custom JAR  
Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata", "pathtooutputbucket  
\  
--configurations https://s3.amazonaws.com/mybucket/myfolder/myConfig.json
```

myConfig.json:

```
[  
  {  
    "Classification": "mapred-site",  
    "Properties": {  
      "mapreduce.application.classpath": "path1,path2"  
    }  
  }  
]
```

パスのカンマ区切りリストを、各タスクの JVM クラスパスに付加する必要があります。

## YARN コンテナの不均一なメモリアクセス対応を有効にする

Amazon EMR バージョン 6.x 以降では、不均一なメモリアクセス (NUMA) を使用してクラスター上のデータをマルチプロセッシングできます。NUMA はプロセッサが別のプロセッサ上のメモリやプロセッサ間で共有されているメモリよりも高速に自身のローカルメモリにアクセスできるコンピュータメモリの設計パターンです。YARN コンテナは、後続のすべてのメモリ割り当てを処理する特定の NUMA ノードにバインドできるため、NUMA でパフォーマンスが向上します。これにより、クラスターがリモートメモリアクセスする回数が減ります。

ワーカーノードマシンがマルチ NUMA ノードの場合は YARN コンテナの NUMA サポートを有効にできます。ワーカーノードがシングル NUMA ノードかマルチ NUMA ノードかを確認するには、次のコマンドを実行します。

```
lscpu | grep -i numa
NUMA node(s): 2
```

一般に、12x を超えるインスタンスには 2 つの NUMA ノードがあります。これは、メタルインスタンスには適用されません。

YARN コンテナの NUMA 対応を有効にするには

1. Amazon EMR 6.x クラスターで次の `yarn-site` 設定を使用します。

```
[
  {
    "classification": "yarn-site",
    "properties": {
      "yarn.nodemanager.linux-container-executor.nonsecure-mode.local-user": "yarn",
      "yarn.nodemanager.linux-container-executor.group": "yarn",
      "yarn.nodemanager.container-executor.class": "org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor",
      "yarn.nodemanager.numa-awareness.enabled": "true",
      "yarn.nodemanager.numa-awareness.numactl.cmd": "/usr/bin/numactl",
      "yarn.nodemanager.numa-awareness.read-topology": "true"
    },
    "configurations": []
  }
]
```

2. クラスターに次のブートストラップアクションを指定します。

```
#!/bin/bash

sudo yum -y install numactl
echo 1 | sudo tee /proc/sys/kernel/numa_balancing

echo "banned.users=mapred,bin,hdfs" >> /etc/hadoop/conf/container-executor.cfg
rm -rf /var/log/hadoop-yarn/
sudo chown -R yarn:hadoop /var/log/hadoop-yarn/
sudo chmod 755 -R /var/log/hadoop-yarn/

sudo chmod 6050 /etc/hadoop/conf/container-executor.cfg

mkdir /mnt/yarn && sudo chmod 755 -R /mnt/yarn && sudo chown -R yarn:hadoop /mnt/
yarn
mkdir /mnt1/yarn && sudo chmod 755 -R /mnt1/yarn && sudo chown -R yarn:hadoop /
mnt1/yarn
mkdir /mnt2/yarn && sudo chmod 755 -R /mnt2/yarn && sudo chown -R yarn:hadoop /
mnt2/yarn
```

- すべてのコンテナが NUMA に対応している必要があります。各コンテナ内の Java 仮想マシン (JVM) に NUMA フラグで通知できます。例えば、JVM に MapReduce ジョブで NUMA を使用するように通知するには、`mapred-site.xml` に次のプロパティを追加します。

```
<property>
  <name>mapreduce.reduce.java.opts</name>
  <value>-XX:+UseNUMA</value>
</property>
<property>
  <name>mapreduce.map.java.opts</name>
  <value>-XX:+UseNUMA</value>
</property>
```

- NUMA が有効になっていることを確認するには、次のコマンドで NodeManager ログファイルを検索します。

```
grep "NUMA resources allocation is enabled," *
```

NodeManager が NUMA ノードリソースをコンテナに割り当てたことを確認するには、次のコマンドで NodeManager ログを検索し、を独自のコンテナ ID `<container_id>` に置き換えます。

```
grep "NUMA node" | grep <container_id>
```

## Hadoop のバージョン履歴

次の表は、Amazon EMR の各リリースに含まれている Hadoop のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Hadoop バージョン情報

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.36.2	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.1.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resour



Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
		cemanager, hadoop-yarn-timeline-server
emr-7.0.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.15.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.14.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.13.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.12.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.11.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.11.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.10.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.10.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.9.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.9.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.8.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.8.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.7.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.36.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.36.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server



Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.6.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.35.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.5.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.4.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.3.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.3.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.2.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.2.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.1.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.1.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-6.0.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.0.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.34.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.33.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.33.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.32.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server



Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.32.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.31.1	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.31.0	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.30.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.30.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.30.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.29.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.28.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.28.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.27.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.27.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.26.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.25.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.24.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.24.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.23.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server



Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.23.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.22.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.21.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.21.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.21.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.20.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.20.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.19.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.19.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.18.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.18.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.17.2	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.17.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.17.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.16.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.16.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server



Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.15.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.15.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.14.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.14.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.14.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.13.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.13.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.3	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.12.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.12.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.11.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.11.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server



Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.10.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.10.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.9.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.8.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.7.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.7.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.6.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.6.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.5.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.5.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.5.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.4.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.4.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager



Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.3.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.2.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.2.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.1.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.1.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.0.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-5.0.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.6	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-4.9.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-4.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-4.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager



Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-4.8.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.4	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.3	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-4.7.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-4.6.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.6.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.5.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-4.4.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.3.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.2.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR リリースラベル	Hadoop のバージョン	Hadoop でインストールされるコンポーネント
emr-4.1.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.0.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

## Hadoop リリースノート (バージョン別)

### [Amazon EMR 6.6.0 - Hadoop リリースノート](#)

### Amazon EMR 6.6.0 - Hadoop リリースノート

#### Amazon EMR 6.6.0 - Hadoop の変更点

型	説明
バグ	BZip2 テキストファイルを読み込む際の重複レコードを修正しました。

型	説明
バックポート	<a href="#">HADOOP-18136</a> : 見つからない .tar ファイルの <code>FileUtils.unTar()</code> 処理の検証
バックポート	<a href="#">HADOOP-17627</a> : branch-3.2 HADOOP-17371、HADOOP-17621、HADOOP-17625 にバックポートして Jetty を 9.4.39 に更新
バックポート	<a href="#">HADOOP-17655</a> : Jetty を 9.4.40 にアップグレード
バックポート	<a href="#">HADOOP-17796</a> : Jetty バージョンを 9.4.43 にアップグレード
バックポート	<a href="#">HADOOP-17661</a> : <code>mvn versions:set</code> が <code>pom.xml</code> の解析に失敗する
バックポート	<a href="#">HADOOP-17236</a> : CVE-2017-18640 を軽減するために SnakeYAML を 1.26 にアップグレード
バックポート	<a href="#">HADOOP-16717</a> : Log GenericsUtil 4 への <code>isLog4jLogger Log4jLoggerAdapter</code>
バックポート	<a href="#">HADOOP-17633</a> : CVE のため <code>json-smart</code> を 2.4.2 <code>nimbus-jose-jwt</code> に、9.8 にバンプする CVEs
バックポート	<a href="#">HADOOP-17844</a> : JSON smart を 2.4.7 にアップグレード
バックポート	<a href="#">HADOOP-17972</a> : branch-3.2 に HADOOP-17683 ( <code>commons-io</code> を 2.8.0 に更新) をバックポート
バックポート	<a href="#">HADOOP-16555</a> : <code>commons-compress</code> を 1.19 に更新

型	説明
バックポート	<a href="#">HADOOP-17370</a> : commons-compress を 1.21 にアップグレード
バックポート	<a href="#">HADOOP-17096</a> : Z StandardCompressor 入力バッファオフセットを修正
バックポート	<a href="#">HADOOP-17112</a> : コミッター経由で s3a にファイルを保存する際、パスに空白が使用できない
バックポート	<a href="#">HADOOP-13500</a> : 設定プロパティオブジェクトのイテレーションの同期
バックポート	<a href="#">HDFS-14099</a> : Z で複数のフレームを解凍するときの不明なフレーム記述子 StandardDe compressor
バックポート	<a href="#">HDFS-16410</a> : での安全でない Xml 解析 OfflineEditsXMLLoader
バックポート	<a href="#">HDFS-14498</a> : 作成に失敗したファイルで永久にループ LeaseManager できる
バックポート	<a href="#">HDFS-15290</a> : 起動 HttpServer 時の NameNode の NPE
バックポート	<a href="#">HDFS-15293</a> : チェックポイントを受信したときの fsimage の受け入れ条件を緩和
バックポート	<a href="#">HDFS-12979</a> : チェックポイント ObserverNode 後に にアップロード StandbyNode Fslmageする必要があります
バックポート	<a href="#">YARN-10538</a> : AM に返却される更新済みノードのリストに再稼働ノードを追加

型	説明
バックポート	<a href="#">YARN-10472</a> : YARN-10314 (YarnClient はシェーディングされたクライアント jar のみ NoClassDefFoundError WebSocketException を持つ をスローします) を branch-3.2 にバックポートします
バックポート	<a href="#">YARN-9968</a> : NullPointerException が原因で Public Localizer が NodeManager で終了する
バックポート	<a href="#">YARN-10651</a> : AbstractYarnScheduler.updateNodeResource() の NPE で CapacityScheduler がクラッシュする
バックポート	<a href="#">YARN-9339</a> : アプリケーションを新しいキューに移動すると、保留中のアプリケーションのメトリクスが正しく表示されない
バックポート	<a href="#">YARN-10438</a> : ClientRMService#getContainerReport() で null containerId を処理
バックポート	<a href="#">YARN-7266</a> : RollingLevelDb ファイルが破損しているか、見つからないと ATS 1.5 が起動しない
バックポート	<a href="#">YARN-9063</a> : RollingLevelDb ファイルが破損しているか、見つからないと ATS 1.5 が起動しない
バックポート	<a href="#">YARN-9848</a> : YARN-4946 を元に戻す (RM はログ集約が終了状態ではない場合、アプリケーションを COMPLETED と見なさないようにする)。



# Apache HBase

[HBase](#) は、Apache Software Foundation の Hadoop プロジェクトの一部として開発されたオープンソースの、非リレーショナルな、分散型データベースです。HBase は、Hadoop Distributed File System (HDFS) 上で実行され、Hadoop エコシステムの非リレーショナルデータベース機能を提供します。HBase は、Amazon EMR リリースバージョン 4.6.0 以降に含まれています。

HBase は Hadoop とシームレスに連携し、ファイルシステムを共有し、MapReduce フレームワークと実行エンジンへの直接入出力として機能します。また、HBase テーブルに対する SQL のようなクエリ、Hive ベースのテーブルとの結合、および Java Database Connectivity ( JDBC ) を有効にすることで、Apache Hive とも統合されます。HBase の詳細については、Apache ウェブサイトで [Apache HBase](#) および [HBase ドキュメント](#) を参照してください。Hive で HBase を使用方法の例については、AWS Big Data Blog の記事「[Buin NoSQL and massively parallel analytics using Apache HBase and Apache Hive on Amazon EMR](#)」を参照してください。

Amazon EMR の HBase では、HBase データを Amazon Simple Storage Service (Amazon S3) に直接バックアップできます。また、HBase クラスターの起動時に以前作成したバックアップから復元できます。Amazon EMR には、データ保持と災害対策の目的で Amazon S3 と統合する追加オプションがあります。

- Amazon S3 の HBase - Amazon EMR バージョン 5.2.0 以降では、Amazon S3 で HBase を使用して、クラスターの HBase ルートディレクトリとメタデータを直接 Amazon S3 に保存できます。その後に起動した新しいクラスターでは、Amazon S3 のルートディレクトリの場所を参照できます。一度に 1 つのクラスターのみが Amazon S3 の HBase の場所を使用できます (リードレプリカクラスターを除く)。詳細については、「[Amazon S3 の HBase \(Amazon S3 ストレージモード\)](#)」を参照してください。
- HBase リードレプリカ - Amazon S3 の HBase で使用する Amazon EMR バージョン 5.7.0 以降では、リードレプリカクラスターがサポートされます。リードレプリカクラスターは、プライマリクラスターのストアファイルおよびメタデータの読み取り専用オペレーションに対して読み取り専用アクセスを提供します。詳細については、「[リードレプリカクラスターの使用](#)」を参照してください。
- HBase スナップショット - Amazon S3 の HBase の代わりに、EMR バージョン 4.0 以降では HBase データのスナップショットを直接 Amazon S3 に作成し、このスナップショットを使用してデータを復元できます。詳細については、「[HBase のスナップショットを使用する](#)」を参照してください。

**⚠ Important**

Amazon EMR HBase クラスタースケールリングでは、[マネージドスケールリング](#)や [HBase クラスターでのカスタムポリシーによるスケールリング](#)の使用はお勧めしません。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている HBase のバージョンと、Amazon EMR で HBase と共にインストールされるコンポーネントを示しています。

このリリースで HBase と共にインストールされるコンポーネントのバージョンについては、[「リリース 7.1.0 コンポーネントバージョン」](#)を参照してください。

## emr-7.1.0 の HBase バージョン情報

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-7.1.0	HBase 2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている HBase のバージョンと、Amazon EMR で HBase と共にインストールされるコンポーネントを示しています。

このリリースで HBase と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

### emr-6.15.0 の HBase バージョン情報

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.15.0	HBase 2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

#### Note

Apache HBase HBCK2 は、HBase のリージョンとシステムテーブルを修復するための別個の運用ツールです。Amazon EMR バージョン 6.1.0 以降では、hbase-hbck2.jar はプライマリノード/usr/lib/hbase-operator-tools/で提供されます。このツールのビルドと使用の詳しい方法については、「[HBase HBCK2](#)」を参照してください。

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている HBase のバージョンと、Amazon EMR で HBase と共にインストールされるコンポーネントを示しています。

このリリースで HBase と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

### emr-5.36.2 の HBase バージョン情報

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.36.2	HBase 1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

### トピック

- [HBase を使用するクラスターの作成](#)
- [Amazon S3 の HBase \(Amazon S3 ストレージモード\)](#)
- [Amazon EMR のログ先行書き込み \(WAL\)](#)
- [HBase シェルを使用する](#)
- [Hive で HBase テーブルにアクセスする](#)
- [HBase のスナップショットを使用する](#)
- [HBase を設定する](#)
- [HBase ユーザーインターフェイスを表示する](#)
- [HBase ログファイルを表示する](#)
- [Ganglia で HBase を監視する](#)

- [HBase の以前のバージョンからの移行](#)
- [HBase リリース履歴](#)

## HBase を使用するクラスターの作成

このセクションの手順では、AWS Management Console とを使用してクラスターを起動する基本について説明します AWS CLI。Amazon EMR クラスターの計画、設定、起動方法の詳細については、「Amazon EMR 管理ガイド」の「[クラスターの計画と設定](#)」を参照してください。

### コンソールを使用して HBase を含むクラスターを作成する

コンソールを使用してクラスターを起動する簡単な手順については、「Amazon EMR 管理ガイド」の「[Amazon EMR の使用開始](#)」を参照してください。

コンソールを使用し、HBase がインストールされた状態でクラスターを作成するには

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Software Configuration (ソフトウェア設定)] で、Amazon リリースバージョン 4.6.0 以降を選択します (最新バージョンをお勧めします)。必要に応じて、HBase などのアプリケーションを選択します。
4. Amazon EMR バージョン 5.2.0 以降の [HBase Storage Settings] (HBase ストレージ設定) で、[HDFS] または [S3] を選択します。詳細については、「[Amazon S3 の HBase \(Amazon S3 ストレージモード\)](#)」を参照してください。
5. 必要に応じて他のオプションを選択し、[Create cluster (クラスターの作成)] を選択します。

### を使用して HBase でクラスターを作成する AWS CLI

HBase がインストールされた状態でクラスターを作成するには、次のコマンドを使用します。

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.1.0 \  
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \  
--instance-type m5.xlarge --instance-count 3
```

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

Amazon S3 で HBase を使用する場合は、JSON 設定オブジェクトを参照する `--configurations` オプションを指定します。設定オブジェクトには、`hbase.rootdir` プロパティを使用して HBase データが格納された Amazon S3 内の場所を指定する `hbase-site` 分類が含まれている必要があります。また、`hbase` プロパティを使用して `s3` を指定する `hbase.emr.storageMode` 分類が含まれている必要があります。以下の例では、これらの設定を使用した JSON スニペットについて説明します。

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "hbase.rootdir": "s3://MyBucket/MyHBaseStore"
    }
  },
  {
    "Classification": "hbase",
    "Properties": {
      "hbase.emr.storageMode": "s3"
    }
  }
]
```

Amazon S3 の HBase の詳細については、「[Amazon S3 の HBase \(Amazon S3 ストレージモード\)](#)」を参照してください。分類の詳細については、[アプリケーションの設定](#) を参照してください。

## Amazon S3 の HBase (Amazon S3 ストレージモード)

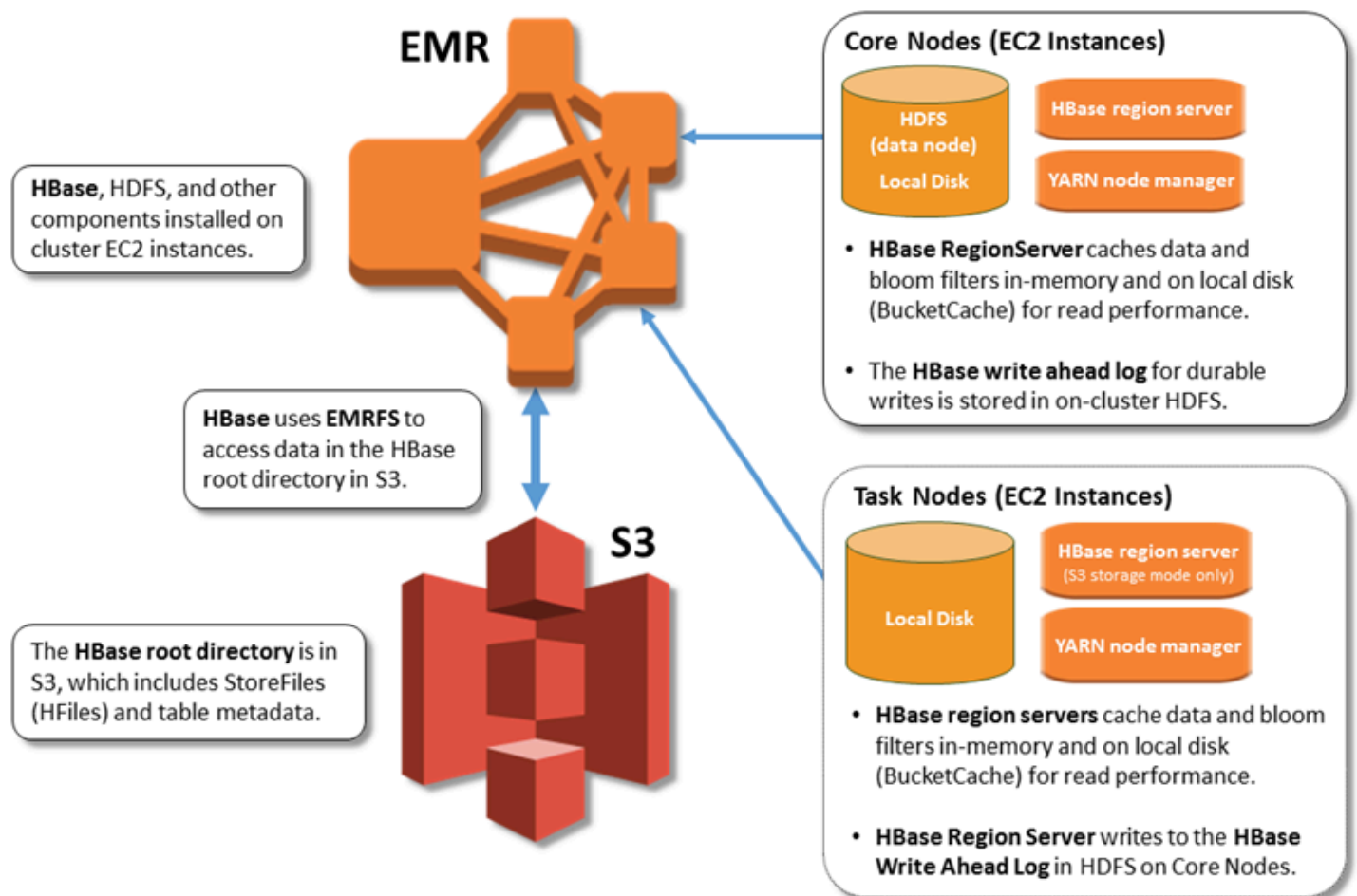
Amazon EMR バージョン 5.2.0 以降で HBase を実行する場合は、Amazon S3 で HBase を有効にすることで、以下の利点を得られます。

- HBase ルートディレクトリが、HBase ストアファイルおよびテーブルメタデータも含めて、Amazon S3 に保存されます。このデータはクラスター外で永続的で、Amazon EC2 アベイラ

ビリティゾーン間で使用することができ、スナップショットやその他の方法を使用して復元する必要はありません。

- Amazon S3 に格納したファイルを使用して、HDFS の 3 倍のレプリケーションで、データ要件ではなくコンピューティング要件に合わせて Amazon EMR クラスターのサイズを決定できます。
- Amazon EMR バージョン 5.7.0 以降では、リードレプリカクラスターを設定して、データの読み取り専用コピーを Amazon S3 に保持できます。複数の読み取りオペレーションを同時に実行する場合や、プライマリクラスターが使用不能になった場合に、リードレプリカクラスターのデータを利用できます。
- Amazon EMR バージョン 6.2.0 以降では、永続的 HFile トラッキングは `hbase:storefile` と呼ばれる HBase システムテーブルを使用し、読み取りオペレーションに使用される HFile パスを直接追跡します。この機能はデフォルトで有効になっており、手動マイグレーションを実行する必要はありません。

次の図は Amazon S3 の HBase に関連する HBase コンポーネントを示しています。





## Amazon S3 での HBase の有効化

Amazon EMR コンソール Amazon S3 で HBase を有効にできます。AWS CLI この設定は、クラスター作成時のオプションです。コンソールを使用する場合は、[Advanced options (詳細オプション)] を使用して設定を選択します。AWS CLIを使用する場合は、`--configurations` オプションを使用して JSON 設定オブジェクトを提供します。設定オブジェクトのプロパティは Amazon S3 内のストレージモードとルートディレクトリの場所を指定します。Amazon S3 の場所として Amazon EMR クラスターと同じリージョン内の場所を指定する必要があります。一度に 1 つのアクティブなクラスターのみが Amazon S3 の同じ HBase ルートディレクトリを使用できます。コンソールの手順と、を使用した詳細な `create-cluster` の例については AWS CLI、「」を参照してください [HBase を使用するクラスターの作成](#)。設定オブジェクトの例は、次の JSON スニペットに示します。

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.rootdir": "s3://my-bucket/my-hbase-rootdir"
  },
}
{
  "Classification": "hbase",
  "Properties": {
    "hbase.emr.storageMode": "s3"
  }
}
```

### Note

Amazon S3 バケットを HBase 用の `rootdir` として使用する場合、Amazon S3 URI の末尾にスラッシュを追加する必要があります。例えば、問題を回避するために、`"hbase.rootdir: s3://my-bucket"` ではなく、`"hbase.rootdir: s3://my-bucket/"` を使用する必要があります。

## リードレプリカクラスターの使用

Amazon S3 の HBase を使用してプライマリクラスターを設定すると、プライマリクラスターと同じデータへの読み取り専用アクセスを提供するリードレプリカクラスターを作成して設定できます。これは、クエリデータに同時にアクセスする場合や、プライマリクラスターの使用不能時にアクセスを



絶やさない場合に便利です。リードレプリカ機能は Amazon EMR バージョン 5.7.0 以降で使用できます。

プライマリクラスターとリードレプリカクラスターは 1 つの重要な点を除いて設定方法が同じです。どちらも同じ `hbase.rootdir` の場所を参照します。ただし、リードレプリカクラスターの `hbase` 分類には `"hbase.emr.readreplica.enabled": "true"` プロパティが含まれます。

たとえば、このトピックで以前に示したプライマリクラスターの JSON 分類の場合、リードレプリカクラスターの設定は次のとおりです。

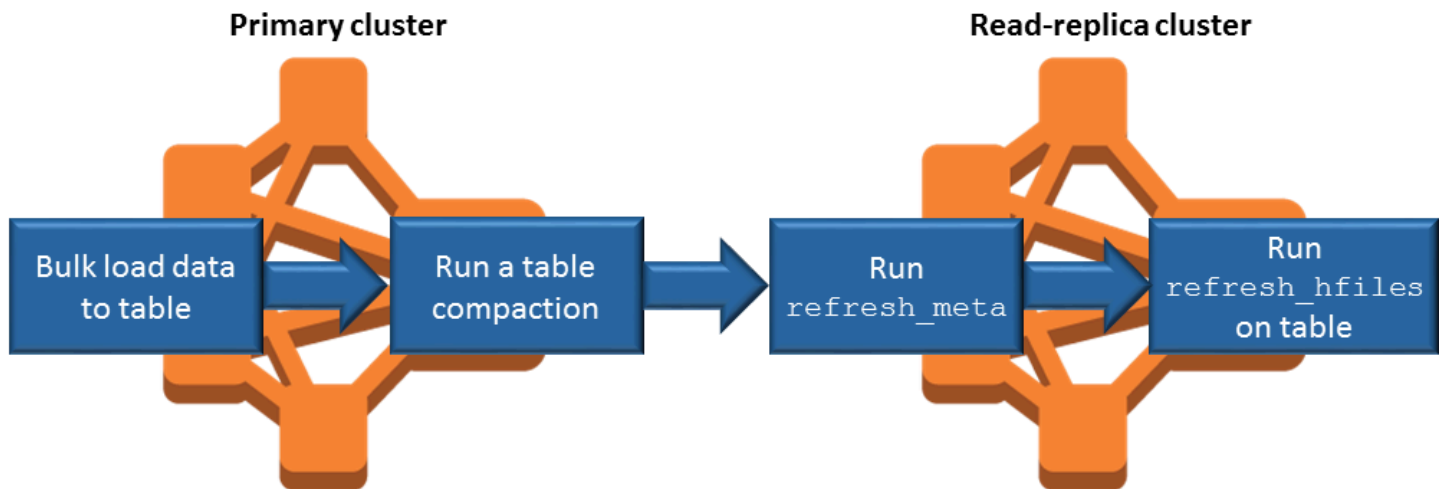
```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.rootdir": "s3://my-bucket/my-hbase-rootdir"
  },
  {
    "Classification": "hbase",
    "Properties": {
      "hbase.emr.storageMode": "s3",
      "hbase.emr.readreplica.enabled": "true"
    }
  }
}
```

## データの追加に伴うリードレプリカの同期化

リードレプリカは、プライマリクラスターが Amazon S3 に書き込む HBase StoreFiles とメタデータを使用するため、リードレプリカは Amazon S3 データストアと同じくらい最新です。Amazon S3 データを書き込むまでにプライマリクラスターとリードレプリカの間で生じる時間差を最小化するには、以下のガイダンスを参考にしてください。

- 可能な限り、プライマリクラスターのデータを一括ロードします。詳細については、Apache HBase ドキュメントの「[一括ロード](#)」を参照してください。
- ストアファイルを Amazon S3 に書き込むフラッシュは、データの追加後、できるだけ早急に発生する必要があります。時間差が最小になるように手動でフラッシュするか、フラッシュ設定を調整します。
- 圧縮が自動実行される場合、圧縮がトリガーされるときの不整合を避けるために手動圧縮を実行します。

- リードレプリカクラスターで、メタデータの変更 (HBase リージョンの分割、圧縮の発生、テーブルの追加や削除など) があった場合、`refresh_meta` コマンドを実行します。
- リードレプリカクラスターで、レコードがテーブルに追加されるかテーブル内で変更された場合、`refresh_hfiles` コマンドを実行します。



## 永続的 HFile トラッキング

永続的 HFile トラッキングでは、`hbase:storefile` と呼ばれる HBase システムテーブルを使用し、読み取りオペレーションに使用される HFile パスを直接追跡します。HBase に追加のデータが追加されると、新しい HFile パスがテーブルに追加されます。これにより、重要な書き込みパス HBase オペレーションのコミットメカニズムとして名前変更オペレーションが削除され、ファイルシステムディレクトリの一覧ではなく、`hbase:storefile` システムテーブルから読み取ることで HBase リージョンを開くときに復旧時間が改善されます。この機能は Amazon EMR バージョン 6.2.0 以降ではデフォルトで有効になっており、手動での移行手順は必要ありません。

### Note

HBase ストアファイルシステムテーブルを使用した永続的な HFile トラッキングは、HBase リージョンのレプリケーション機能をサポートしていません。HBase リージョンレプリケーションの詳細については、「[Timeline-consistent high available reads](#)」を参照してください。

## 永続的 HFile トラッキングの無効化

Amazon EMR リリース 6.2.0 以降、永続的な HFile 追跡はデフォルトで有効になっています。永続的 HFile トラッキングを無効にするには、クラスターの起動時に次の設定オーバーライドを指定します。

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.storefile.tracking.persist.enabled": "false",

    "hbase.hstore.engine.class": "org.apache.hadoop.hbase.regionserver.DefaultStoreEngine"
  }
}
```

#### Note

Amazon EMR クラスターを再設定するときは、すべてのインスタンスグループを更新する必要があります。

## ストアファイルテーブルの手動同期

新しい HFiles が作成されると、ストアファイルテーブルは最新の状態に保たれます。ただし、何らかの理由でストアファイルテーブルがデータファイルと同期しなくなった場合は、次のコマンドを使用して手動でデータを同期できます。

オンラインリージョンでストアファイルテーブルを同期する:

```
hbase org.apache.hadoop.hbase.client.example.RefreshHFilesClient <table>
```

オフラインリージョンでストアファイルテーブルを同期する:

- ストアファイルテーブル znode を削除します。

```
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
[<tableName>, hbase:namespace]
# The TableName exists in the list
echo "delete /hbase/storefile/loaded/<tableName>" | sudo -u hbase hbase zkcli
# Delete the Table ZNode
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
[hbase:namespace]
```

- リージョンを割り当てます ('hbase shell' で実行します)。

```
hbase cli> assign '<region name>'
```

- 割り当てが失敗した場合。

```
hbase cli> disable '<table name>'  
hbase cli> enable '<table name>'
```

## ストアファイルテーブルのスケーリング

ストアファイルテーブルは、デフォルトでは 4 つのリージョンに分割されます。ストアファイルテーブルの書き込み負荷が大きい場合は、テーブルを手動で分割できます。

特定のホットリージョンを分割するには、次のコマンドを使用します ('hbase shell' で実行します)。

```
hbase cli> split '<region name>'
```

テーブルを分割するには、次のコマンドを使用します ('hbase shell' で実行します)。

```
hbase cli> split 'hbase:storefile'
```

## 運用上の考慮事項

HBase リージョンサーバーは BlockCache、 を使用してデータ読み取りをメモリに保存し、データ読み取り BucketCache をローカルディスクに保存します。さらに、リージョンサーバーは MemStore を使用してデータをメモリ内に書き込み、先行書き込みログを使用してデータを HDFS に保存してから、データを Amazon S3 の HBase StoreFiles に書き込みます。クラスターの読み込みパフォーマンスは、レコードがどのくらいの頻度でインメモリキャッシュまたはオンディスクキャッシュから取得できるかに関連しています。キャッシュミスが発生すると、レコードは Amazon S3 StoreFile の から読み取られます。これにより、HDFS から読み取るよりもレイテンシーが大きくなり、標準偏差も高くなります。Amazon S3 さらに、Amazon S3 の最大要求率は、ローカルキャッシュから受理できるものよりも低くなるため、データのキャッシュは読み取り過多のワークロードにとって重要である場合があります。Amazon S3 のパフォーマンスの詳細については、「Amazon Simple Storage Service ユーザーガイド」の「[パフォーマンスの最適化](#)」を参照してください。

パフォーマンスを向上させるには、EC2 インスタンスストレージに可能な限りデータセットをキャッシュすることをお勧めします。はリージョンサーバーの EC2 インスタンススト

レージ BucketCache を使用するため、十分なインスタンスストアを持つ EC2 インスタンスタイプを選択し、必要なキャッシュサイズに合わせて Amazon EBS ストレージを追加できます。hbase.bucketcache.size プロパティを使用して、アタッチされたインスタンスストアと EBS ボリューム BucketCache のサイズを増やすこともできます。デフォルトの設定は 8,192 MB です。

書き込みの場合、MemStore フラッシュの頻度と、マイナー圧縮とメジャー圧縮中 StoreFiles に存在する の数は、リージョンサーバーの応答時間の増加に大きく影響する可能性があります。最適なパフォーマンスを得るには、MemStore フラッシュブロックと HRegion ブロックの乗数のサイズを増やすことを検討してください。これにより、メジャー圧縮間の経過時間が長くなりますが、リードレプリカを使用すると整合性のある遅延も長くなります。場合によっては、ファイルブロックサイズを増やして (ただし 5 GB 未満)、EMRFS で Amazon S3 マルチパートアップロード機能をトリガーすることで、パフォーマンスを向上できる場合があります。Amazon EMR のブロックサイズのデフォルトは 128 MB です。詳細については、「[HDFS 構成](#)」を参照してください。フラッシュと圧縮に関するパフォーマンスのベンチマークで 1 GB のブロックサイズを超えているお客様はほとんどいません。さらに、HBase 圧縮とリージョンサーバーは、圧縮 StoreFiles する必要性が少ない場合に最適に動作します。

大きなディレクトリの名前を変更する必要があるため、Amazon S3 でのテーブルの削除に長い時間がかかる場合があります。テーブルを削除する代わりに無効化することを検討してください。

古い WAL ファイルとストアファイルをクリーンアップする HBase クリーナープロセスがあります。Amazon EMR リリースバージョン 5.17.0 以降では、クリーナーがグローバルに有効になっており、次の設定プロパティをクリーナーの動作の制御に使用できます。

設定プロパティ	デフォルト値	説明
hbase.regionserver.hfilecleaner.large.thread.count	1	期限が切れた大規模な HFiles をクリーンアップするために割り当てられたスレッドの数。
hbase.regionserver.hfilecleaner.small.thread.count	1	期限が切れた小規模な HFiles をクリーンアップするために割り当てられたスレッドの数。

設定プロパティ	デフォルト値	説明
<code>hbase.cleaner.scan.dir.concurrent.size</code>	すべての使用可能なコアの 4 分の 1 に設定します。	oldWALs ディレクトリをスキャンするスレッドの数。
<code>hbase.oldwals.cleaner.thread.size</code>	2	oldWALs ディレクトリの WALs をクリーンアップするスレッドの数。

Amazon EMR 5.17.0 以前では、高負荷のワークロードを実行しているときにクリーナーオペレーションがクエリパフォーマンスに影響する可能性があるため、ピークを過ぎた時間帯だけクリーナーを有効にすることをお勧めします。クリーナーには、次の HBase シェルコマンドがあります。

- `cleaner_chore_enabled` クリーナーが有効かどうかのクエリ。
- `cleaner_chore_run` 手動でクリーナーを実行してファイルを削除します。
- `cleaner_chore_switch` クリーナーを有効、または無効にし、クリーナーの以前の状態に戻します。たとえば、`cleaner_chore_switch true` はクリーナーを有効にします。

## Amazon S3 の HBase のパフォーマンスをチューニングするプロパティ

Amazon S3 の HBase を使用するとき、以下のパラメータを調整してワークロードのパフォーマンスをチューニングできます。

設定プロパティ	デフォルト値	説明
<code>hbase.bucketcache.size</code>	8,192	リージョンサーバー Amazon EC2 インスタンスストアと BucketCache EBS ボリュームでストレージ用に予約されているディスク容量を MB 単位で表したものです。設定は、すべてのリージョンサーバーインスタンスに適用されません。通常、BucketCache サイズ

設定プロパティ	デフォルト値	説明
		ズが大きいほどパフォーマンスが向上します。
hbase.hregion.memstore.flush.size	134217728	Amazon S3 への Memstore フラッシュがトリガーされるデータ上限 (バイト単位)。
hbase.hregion.memstore.block.multiplier	4	更新がブロックされる MemStore 上限を決定する乗数。がこの値に乗算された hbase.hregion.memstore.flush.size MemStore を超えると、更新はブロックされます。MemStore flushes が発生し、圧縮によって更新のブロックが解除される可能性があります。
hbase.hstore.blockingStoreFiles	10	更新 StoreFiles がブロックされる前にストアに存在することができるの最大数。
hbase.hregion.max.filesize	10737418240	リージョンが分割されるまでのリージョンの最大サイズ。

## データ損失のないクラスターのシャットダウンと復元

Amazon S3 に書き込まれていないデータを失うことなく Amazon EMR クラスターをシャットダウンするには、MemStore キャッシュを Amazon S3 にフラッシュして新しいストアファイルを書き込む必要があります。まず、すべてのテーブルを無効にする必要があります。次の設定ステップはクラスターにステップを追加する際に使用することができます。詳細については「Amazon EMR 管理ガイド」の「[AWS CLI およびコンソールを使用した手順の作業](#)」を参照してください。

```
Name="Disable all tables",Jar="command-runner.jar",Args=["/bin/bash","/usr/lib/hbase/bin/disable_all_tables.sh"]
```

または、次の Bash コマンドを直接実行できます。

```
bash /usr/lib/hbase/bin/disable_all_tables.sh
```

すべてのテーブルを無効にした後に、HBase シェルと次のコマンドを使用して hbase:meta テーブルをフラッシュします。

```
flush 'hbase:meta'
```

次に、Amazon EMR クラスターで提供されているシェルスクリプトを実行して、MemStore キャッシュをフラッシュできます。これをステップとして追加するか、またはクラスター上の AWS CLI を使用して直接実行できます。このスクリプトは、すべての HBase テーブルを無効にします。これにより、各リージョンサーバーの MemStore Amazon S3 にフラッシュされます。スクリプトが正常に完了すると、データは Amazon S3 に維持され、クラスターを終了できます。

同じ HBase データを持つクラスターを再起動するには、AWS Management Console または hbase.rootdir 設定プロパティを使用して、前のクラスターと同じ Amazon S3 の場所を指定します。

## Amazon EMR のログ先行書き込み (WAL)

Amazon EMR 6.15 以降では、Apache HBase 先行書き込みログ (WAL) を Amazon EMR WAL に書き込むことができます。Amazon EMR のリリースが低い場合、Amazon S3 で HBase Amazon S3 オプションを使用してクラスターを作成すると、WAL はクラスターのローカルディスクに保存される唯一の Apache HBase コンポーネントであり、ルートディレクトリ、ファイル (HFiles)、テーブルメタデータ、Amazon S3 上のデータなどの他のコンポーネントを保存できます。

Amazon EMR WAL を使用して、Amazon S3 にフラッシュされなかったデータを復元できます。HBase クラスターを完全にバックアップするには、Amazon EMR WAL サービスの使用をオプトインします。バックグラウンドで、HBase 先行書き込みログ (WAL) を Amazon EMR の WAL に RegionServer 書き込みます。

クラスターまたは AZ が異常または使用できなくなった場合は、新しいクラスターを作成し、同じ S3 ルートディレクトリと Amazon EMR WAL ワークスペースにポイントして、数分以内に WAL でデータを自動的に復元できます。詳細については、「[Amazon EMR WAL からの復元](#)」を参照してください。



### Note

Amazon EMR は、先行書き込みログとそのデータを、クラスターの作成から 30 日間保持します。30 日後、Amazon EMR は Amazon EMR WAL とそのデータを自動的に削除します。ただし、同じ S3 ルートディレクトリから新しい WAL 対応クラスターを起動する場合、新しいクラスターの起動時刻から 30 日間 WAL の使用を延長できます。Amazon EMR は、最初の 30 日間が経過しても、最初のクラスターからの WAL データをクリーンアップします。詳細については、「[Amazon EMR WAL からの復元](#)」を参照してください。

以下のセクションでは、HBase 対応 EMR クラスターで Amazon EMR WAL をセットアップして使用する方法について説明します。

### トピック

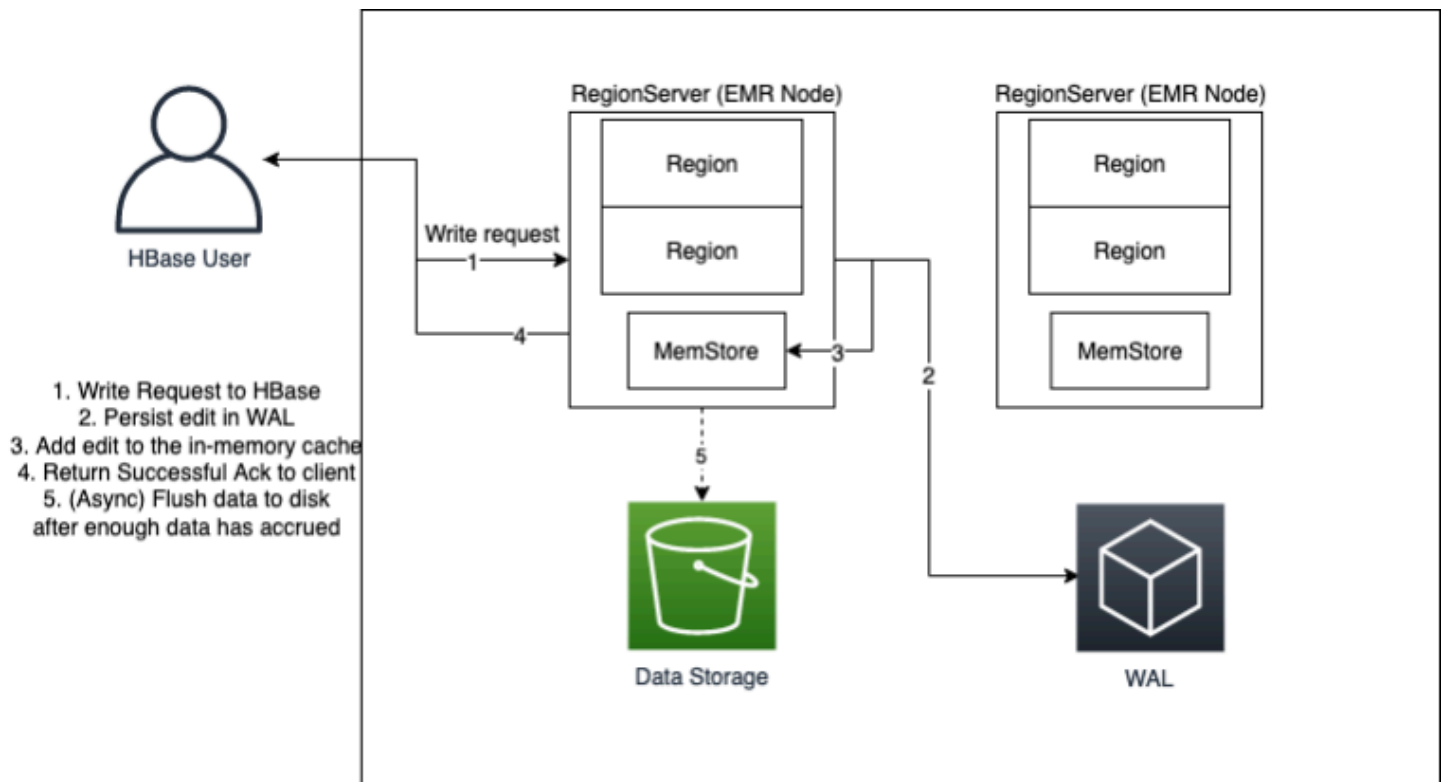
- [Amazon EMR WAL ワークスペース](#)
- [Amazon EMR WAL に必要なアクセス許可](#)
- [Amazon EMR WAL の有効化](#)
- [Amazon EMR WAL からの復元](#)
- [Amazon EMR WAL でのセキュリティ設定の使用](#)
- [経由で Amazon EMR WAL にアクセスする AWS PrivateLink](#)
- [Amazon EMR WAL の料金とメトリクスについて](#)
- [WAL ワークスペースのタグ付け](#)
- [Amazon EMR WAL に関する考慮事項とリージョン](#)
- [Amazon EMR WAL \(EMRWAL\) CLI リファレンス](#)

## Amazon EMR WAL ワークスペース

Amazon EMR WAL に WAL ワークスペースの概念が追加されました。WAL ワークスペースは WAL のWALs。Amazon EMR WAL の各先行書き込みログは、WAL ワークスペースによってカプセル化されます。EMR クラスターは、クラスターの起動時に設定した 1 つの WALs ワークスペースに WAL を書き込みます。ワークスペースを指定defaultWALworkspaceしない場合は、に書き込みます。WAL ワークスペースは、名前空間などの既存の HBase 用語とは関係ありません。

WAL ワークスペースを使用して Amazon EMR WAL IAM アクセス許可の範囲を絞り込み、クラスターがアクセスする必要があるワークスペースのみを含めることができます。タグベースのアクセス

コントロールのために WAL ワークスペースにタグを付けることもできます。タグ付けの詳細については、「」を参照してください[WAL ワークスペースのタグ付け](#)。



## Amazon EMR WAL に必要なアクセス許可

クラスターが Amazon EMR WAL に接続するには、クラスターのインスタンスプロファイルに特定の IAM アクセス許可が必要です。

- Amazon EMR WAL は、[AWSServiceRoleForEMRWAL](#) サービスにリンクされたロールを使用してクラスターのステータスを取得します。Amazon EMR は、WAL ワークスペースを作成するときにこのサービスにリンクされたロールを自動的に作成します。そうしないと、Amazon EMR WAL のワークスペースを設定し、サービスにリンクされたロールがまだ存在しないときに、HBase がサービスにリンクされたロールを作成します。

クラスターで Amazon EMR WAL を有効にする前に、[AWSServiceRoleForEMRWAL](#) サービスにリンクされたロールの自動作成を許可するアクセス許可を設定する必要があります。この機能を追加する詳細とステートメントの例については、「[先行書き込みログ記録にサービスにリンクされたロールを使用する](#)」を参照してください。

- Amazon EMR WAL は HBase Write Ahead Log (WAL) を使用するため、クラスターは HBase WAL を使用する必要があります。HBase を実行するために必要な最小限の IAM アクセス許可を次に示します。インスタンスプロファイルのアクセス許可ポリシーにこれらを追加します。

```
emrwal:DeleteWal
emrwal:CreateWal
emrwal:CreateWorkspace
emrwal:AppendEdit
emrwal:ReplayEdits
emrwal:GetCurrentWalTime
emrwal:CompleteWalFlush
```

### Note

Amazon EMR WAL のアクセス許可を最小セットのみにスコープする場合、一部の [EMRWAL CLI](#) コマンドには実行に必要なアクセス許可がありません。

## Amazon EMR WAL の有効化

を使用してクラスターを作成するときに Amazon EMR WAL への書き込みを有効にするには、次の手順を実行します AWS Command Line Interface。

### Note

すでに実行中のクラスターに対して Amazon EMR WAL を有効にすることはできません。また、同じ S3 ルートディレクトリで 2 つのクラスターを起動することはできません。詳細については、「[Amazon EMR WAL に関する考慮事項とリージョン](#)」を参照してください。

1. Amazon EMR WAL 対応クラスターを作成する前に、クラスターで使用する予定のインスタンスプロファイルに必要なアクセス許可を追加する必要があります。詳細については、「[Amazon EMR WAL に必要なアクセス許可](#)」を参照してください。
2. からクラスターを作成します AWS CLI。以下の例に示すように、`--configurations` オプションを使用して、`hbase.emr.wal.enabled` プロパティを指定する JSON 設定オブジェクトを指定します。
  - Amazon S3 のストレージモードとルートディレクトリの場所を指定します。指定する Amazon S3 の場所は、EMR クラスターと同じ リージョンにある必要がありますが、一度に S3 で同じ HBase ルートディレクトリを使用できるアクティブなクラスターは 1 つだけです。

- インスタンスグループ設定を使用してクラスターを作成します。インスタンスフリート設定で Amazon EMR WAL を使用することはできません。インスタンスグループを使用してクラスターを作成する方法の詳細については、「Amazon EMR [管理ガイド](#)」の「[ユニフォームインスタンスグループを設定する](#)」を参照してください。
  - クラスターを作成するコンソールの手順と、を使用する詳細な create-cluster 例については、[HBase を使用したクラスターの作成 AWS CLI](#) を参照してください。
3. 新しいクラスターの WAL を有効にするには、`hbase.emr.wal.enabled` プロパティを に設定します `true`。次のコマンドには、設定オブジェクトの例を含む JSON スニペットが含まれています。

```
aws emr create-cluster --name "hbasewal" --release-label emr-6.x.y \  
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \  
--instance-type m6i.xlarge --instance-count 1 --configurations hbase.json  
$cat hbase.json  
[  
  {  
    "Classification": "hbase-site",  
    "Properties": {  
      "hbase.rootdir": "s3://MyBucket/MyHBaseStore"  
    }  
  },  
  {  
    "Classification": "hbase",  
    "Properties": {  
      "hbase.emr.storageMode": "s3",  
      "hbase.emr.wal.enabled": "true"  
    }  
  }  
]
```

新しく作成されたクラスターで HBase がオンラインの場合、HBase は自動的に WAL データを Amazon EMR WAL に書き込み、復旧の目的で Amazon EMR WAL を使用します。

#### Example 1: Amazon EMR WAL を使用する EMR クラスターの作成

```
[  
  {  
    "Classification": "hbase-site",  
    "Properties": {  
      "hbase.rootdir": "s3://MyBucket/MyHBaseStore"  
    }  
  }  
]
```

```
    }
  },
  {
    "Classification": "hbase",
    "Properties": {
      "hbase.emr.storageMode": "s3",
      "hbase.emr.wal.enabled": "true"
    }
  }
]
```

## Example 2: カスタム WAL ワークスペースを使用した EMR クラスターの作成

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "hbase.rootdir": "s3://MyBucket/MyHBaseStore",
      "emr.wal.workspace": "customWorkspaceName"
    }
  },
  {
    "Classification": "hbase",
    "Properties": {
      "hbase.emr.storageMode": "s3",
      "hbase.emr.wal.enabled": "true"
    }
  }
]
```

## Amazon EMR WAL からの復元

元のクラスターの Amazon EMR WAL は 30 日間保持されるため、その 30 日以内に新しく作成されたクラスターの WAL を復元して再利用できます。同じ S3 ルートディレクトリから新しいクラスターを起動すると、前の 30 日間の期間が終了していない限り、30 日間のクロックは新しいクラスターの起動時刻から再起動します。

新しいクラスターで既存の WAL を復元するには、次の手順に従います。このプロセスでは、Amazon EMR WAL を有効にして元のクラスターを作成したことを前提としています。

1. WAL 対応クラスターを作成してから 30 日以内に、元のクラスター AWS リージョン と同じ に新しいクラスターを作成します。新しいクラスターは、元のクラスターが作成されたのと同じリージョン内の同じ AZ 内または別の AZ 内に配置できます。

オブジェクトプロパティを設定して、Amazon S3 のストレージモードとルートディレクトリの場所を指定します。指定する Amazon S3 の場所は、EMR クラスターと同じリージョンにある必要がありますが、一度に S3 で同じ HBase ルートディレクトリを使用できるアクティブなクラスターは 1 つだけです。

クラスターを作成するコンソールの手順と、 を使用する詳細なcreate-cluster例については、[HBase を使用したクラスターの作成 AWS CLI](#)」を参照してください。

2. 新しいクラスターに既存の Amazon EMR WAL を使用するには、hbase.emr.wal.enabledプロパティを に設定しますtrue。次の JSON スニペットは、設定オブジェクトの例を示しています。

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "hbase.rootdir": "s3://MyBucket/MyHBaseStore"
    }
  },
  {
    "Classification": "hbase",
    "Properties": {
      "hbase.emr.storageMode": "s3",
      "hbase.emr.wal.enabled": "true"
    }
  }
]
```

## Amazon EMR WAL でのセキュリティ設定の使用

Amazon EMR は、クラスターと Amazon EMR WAL サービス間で転送中のデータと、Amazon EMR WAL に保管中のデータの両方を自動的に暗号化します。詳細については、[「Amazon EMR WAL の保管時の暗号化」](#)を参照してください。セキュリティ設定を使用して、AWS Key Management Service (KMS) サービスから独自のキーを持ち込んで、Amazon EMR WAL に保存するデータを暗号化することもできます。

クラスターの作成時に、次のいずれかの方法を使用してセキュリティ設定を選択します。

## Console

から AWS Management Console、セキュリティ設定 と EC2 キーペア で設定を指定します。

### Security configuration and EC2 key pair - *optional* [Info](#)

#### Security configuration

Select your cluster encryption, authentication, authorization, and instance metadata service settings.

## CLI

から AWS CLI、[create-cluster](#) コマンドを使用するときに `--security-configuration` パラメータを設定します。

詳細については、「[Amazon EMR 管理ガイド](#)」の「[Amazon EMR WAL の保管時の暗号化](#)」および「[セキュリティ設定を使用してクラスターセキュリティを設定する](#)」を参照してください。

WAL に関するセキュリティ関連情報については、「[先行書き込みログのサービスにリンクされたロールの使用](#)」を参照してください。

## 経由で Amazon EMR WAL にアクセスする AWS PrivateLink

AWS ネットワーク内で接続を維持する場合、Amazon EMR WAL は AWS PrivateLink サポートを提供します。を設定するには AWS PrivateLink、AWS Management Console または AWS Command Line Interface (AWS CLI) を使用して、Amazon EMR WAL に接続するインターフェイス VPC エンドポイントを作成します。詳細については、「[AWS PrivateLink ガイド](#)」の「[インターフェイス VPC エンドポイントを使用して AWS サービスにアクセスする](#)」を参照してください。

基本的なステップは次のとおりです。

1. Amazon VPC コンソールを使用して [VPC エンドポイントを作成します](#)。エンドポイントを選択し、エンドポイントを作成します。
2. サービスカテゴリをAWS サービスとして保持します。
3. サービスパネルの検索バーで、と入力し `emrwal`、というラベルの付いたサービスを選択します `com.amazonaws.region.emrwal.prod`。
4. VPC を選択し、エンドポイントを保存します。EMR クラスターにアタッチする VPC エンドポイントに、同じセキュリティグループをアタッチするようにしてください。

5. 必要に応じて、新しいエンドポイントのプライベート DNS ホスト名を有効にできるようになりました。VPC の DNS ホスト名を有効にすると DNS サポートを有効にする `true`を設定します。次に、エンドポイント ID を選択し、アクションメニューから VPC 設定の編集を選択し、プライベート DNS 名を有効にします。
  - エンドポイントのプライベート DNS ホスト名は、 の形式に従います `prod.emrwal.region.amazonaws.com`。
  - プライベート DNS ホスト名を有効にしない場合、Amazon VPC は の形式で DNS エンドポイント名を提供します `endpointID.prod.emrwal.region.vpce.amazonaws.com`。
6. AWS PrivateLink エンドポイントを使用するには、次の例に示すように、[Amazon EMR WAL 対応クラスター](#)の作成時に `emr.wal.client.endpoint`設定を変更します。

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "hbase.rootdir": "s3://MyBucket/MyHBaseStore",
      "emr.wal.workspace": "customWorkspaceName",
      "emr.wal.client.endpoint": "https://prod.emrwal.region.amazonaws.com"
    }
  },
  {
    "Classification": "hbase",
    "Properties": {
      "hbase.emr.storageMode": "s3",
      "hbase.emr.wal.enabled": "true"
    }
  }
]
```

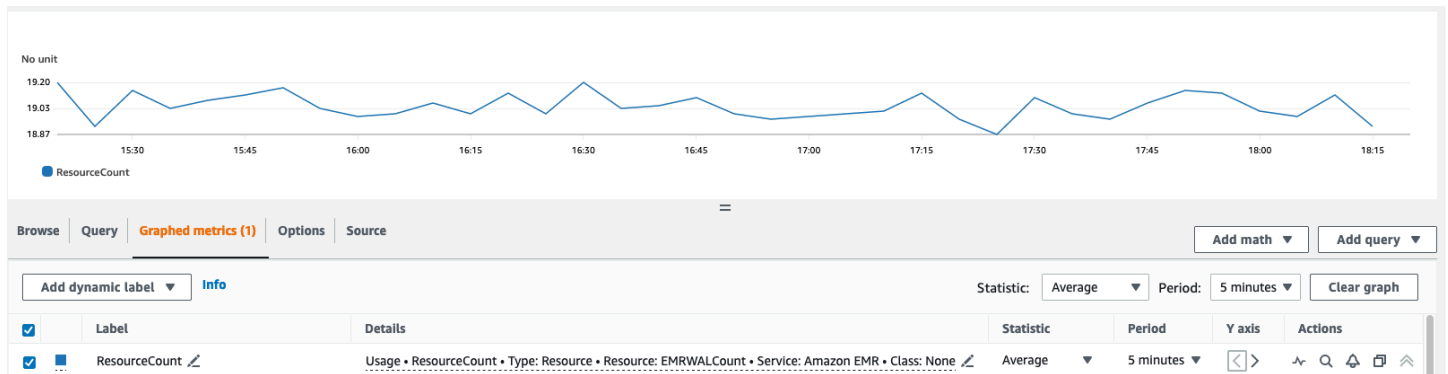
VPCE ポリシーを使用して、Amazon EMR WAL APIs へのアクセスを許可または制限することもできます。詳細については、『AWS PrivateLink ガイド』の「[Control access to VPC endpoints using endpoint policies \(エンドポイントポリシーを使用して VPC エンドポイントへのアクセスをコントロールする\)](#)」を参照してください。



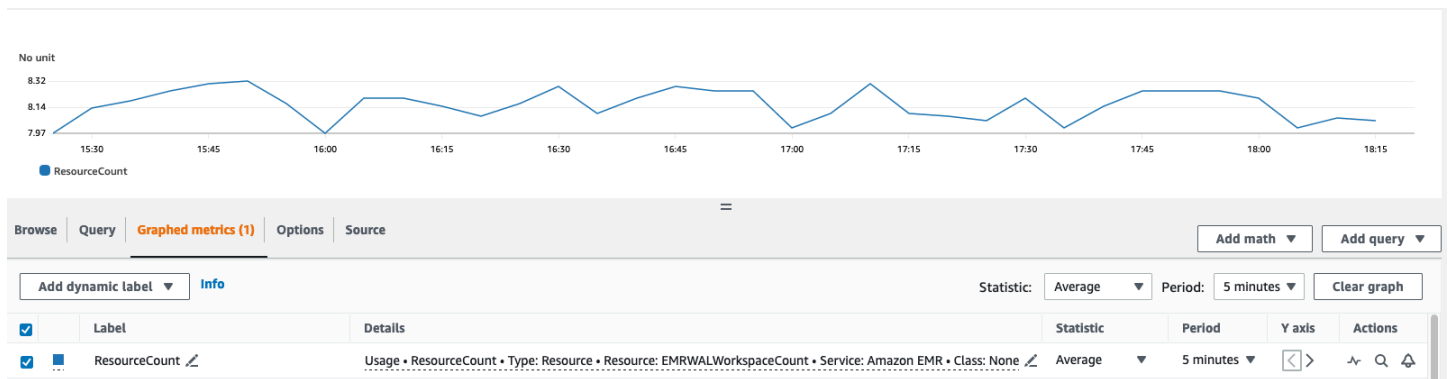
## Amazon EMR WAL の料金とメトリクスについて

コア機能の請求単位	詳細
EMR-WAL-Read-GiB	<p>API calls to read data from your table are billed as ReadRequestGiB. This includes <a href="#">Get and Scan</a> operations. Reads are charged based on the sizes of the read items. Amazon EMR bills at a minimum of 1 byte. For example, if you read a 1234.12 bytes item, you're charged for 1235 bytes. Reads are aggregated every hour for billing and shown as GiBs.</p>
EMR-WAL-Write-GiB	<p>API calls to write data from your table are billed as Write-GiB. This includes <a href="#">Put</a> operations. Writes are charged based on the sizes of the written items. Amazon EMR bills at a minimum of 1 byte. For example, if you write a 1234.12 bytes item, you're charged for 1235 bytes. Writes are aggregated every hour for billing and shown as GiBs.</p>
EMR-WAL-WALHours	<p>サービスに保存する WALs の数は、として請求されます EMR-WAL-WALHours 。 Amazon EMR は HBase リージョンごとに 1 つの WAL を作成します。例えば、システムテーブルを含む 20 個の HBase テーブルを作成し、各テーブルに 2 つの HBase リージョンがある場合、次のように計算された 28,800 WAL 時間を使用します。</p> <div style="border: 1px solid #ccc; border-radius: 10px; padding: 10px; margin-top: 10px;"> <pre> 20 tables x 2 Regions per table x 1 WAL per Region x 30 days x 24 hours ----- 28,800 EMR-WAL-WALHours </pre> </div>

### 例EMRWALCount:



## 例EMRWALWorkspaceCount:



## WAL ワークスペースのタグ付け

新しいワークスペースを作成するときワークスペースにタグを追加したり、実行中のクラスターのアクティブなワークスペースからタグを追加、削除、または一覧表示したりできます。ワークスペース内の個々のリソースにタグを付けることはできず、既存のタグを更新することもできません。代わりに、ワークスペースから不要なタグを削除して置き換えます。

ワークスペースには からタグ付けできますEMRWAL CLI。ワークスペースにタグ付けするためのEMRWAL CLIコマンドのリストについては、「」を参照してください[Amazon EMR WAL \(EMRWAL\) CLI リファレンス](#)。

次の IAM ポリシーの例は、適切なタグ付けキー `resource_tag_allow_test_key` と値を持つワークスペース CRUDL オペレーションのみを許可するシナリオを示しています `resource_tag_allow_test_value`。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
```

```
"Action": [
  "EMRWAL:*"
],
"Effect": "Allow",
"Resource": [
  "*",
  "*"
],
"Condition": {
  "StringEquals": {
    "aws:ResourceTag/resource_tag_allow_test_key": [
      "resource_tag_allow_test_value"
    ]
  }
}
]
```

タグがワークスペースオペレーションに必要なことを確認するには、[Amazon EMR WAL \(EMRWAL\) CLI リファレンス](#) を使用して、目的のリソースタグを持つワークスペース `tagAllowResourceTag` の `listTagsForResource` コマンドを呼び出します。条件を正しく設定すると、コマンドは成功します。

```
emrwal listTagsForResource -r us-east-1 -arn arn:aws:emrwal:us-east-1:arn:workspace/
tagAllowResourceTag
Tag(Key=resource_tag_allow_test_key, Value=resource_tag_allow_test_value)
```

## Amazon EMR WAL に関する考慮事項とリージョン

### Amazon EMR WAL に関する考慮事項

次のリストでは、Amazon EMR WAL の重要な考慮事項と制限事項について説明します。

- Amazon EMR WAL は、Amazon EMR リリース 6.15.0 以降で使用できます。
- Amazon EMR WAL はオプトイン型の有料サービスです。読み取り、書き込み、データストレージの使用に対して料金が発生します。詳細については、[Amazon EMR WAL の料金とメトリクスについて](#)「」および「[Amazon EMR の料金](#)」ページを参照してください。
- Amazon EMR WAL は HBase Write Ahead Log (WAL) を使用します。Amazon EMR WAL を使用するには、クラスターで HBase WAL を使用する必要があります。

- クラスターの作成時に Amazon EMR WAL を有効にするには、必要なロールのアクセス許可が必要です。詳細については、[「先行書き込みログ記録にサービスにリンクされたロールを使用する」](#)を参照してください。
- AWS Management Console、AWS CLI または API を使用してクラスターを作成するときに Amazon EMR WAL を有効にし、インスタンスグループ設定を使用する必要があります。Amazon EMR WAL を使用してクラスターを作成しなかった場合、実行中のクラスターで Amazon EMR WAL を有効にすることはできません。また、実行中のクラスターで Amazon EMR WAL を有効にする hbase-site 設定を編集することはできません。
- Amazon EMR WAL は、ルートディレクトリに Amazon S3 を使用するクラスターでのみ有効にできます。
- Amazon S3 の同じ HBase ルートディレクトリに複数のアクティブなクラスターを持つことはできません。
- リードレプリカクラスターで Amazon EMR WAL を有効にすることはできません。
- WAL は、マネージドサービス内のアベイラビリティーゾーン間でレプリケートされます。
- WAL はクラスターを使い果たし、次のクラスターで引き続き使用できます。
- 起動中またはクラスターが動作中 (実行中の状態) に Amazon EMR WAL を無効にすることはできません。
- WAL とワークスペースの制限については、[「Amazon EMR エンドポイントとクォータ」](#)を参照してください。

## Amazon EMR WAL のリージョンの可用性

Amazon EMR WAL サービスは、次の で利用できます AWS リージョン。

- ap-northeast-1 – アジアパシフィック (東京)
- ap-southeast-1 – アジアパシフィック (シンガポール)
- ap-south-1 – アジアパシフィック (ムンバイ)
- ap-southeast-2 – アジアパシフィック (シドニー)
- eu-central-1 – 欧州 (フランクフルト)
- eu-north-1 – 欧州 (ストックホルム)
- eu-west-1 – 欧州 (アイルランド)
- us-east-1 – 米国東部 (バージニア北部)
- us-east-2 – 米国東部 (オハイオ)

- us-west-2 — 米国西部 (オレゴン)
- ap-northeast-2 – アジアパシフィック (ソウル)

## Amazon EMR WAL (EMRWAL) CLI リファレンス

EMRWAL コマンドラインインターフェイス (EMRWAL CLI) は、Amazon EMR のログ先行書き込み (WAL) を管理するための統合ツールです。クラスターの作成時に WAL を有効にすると、には EMR クラスターがEMRWAL CLI付属しています。WAL の有効化の詳細については、「」を参照してください[Amazon EMR のログ先行書き込み \(WAL\)](#)。

EMRWAL CLI には、次のコマンドが含まれています。

### トピック

- [createWorkspace](#)
- [deleteWal](#)
- [deleteWorkspace](#)
- [listTagsForResource](#)
- [listWals](#)
- [listWorkspaces](#)
- [tagResource](#)
- [untagResource](#)

### createWorkspace

createWorkspace コマンドは、新しい Amazon EMR WAL ワークスペースを作成します。

使用:

```
emrwal createWorkspace [-tags <tags>] [-e {endpoint}] [-r {Region}] -w {workspacename} [-h]
```

例:

```
aws emr createWorkspace -w examples
```

## deleteWal

deleteWals コマンドは、指定した Amazon EMR WAL を削除します。

使用:

```
emrwal deleteWal [-e {endpoint}] [-r {Region}] [-w {workspacename}] [-p <tablePrefix>]
[-n <walName>] [-N <fullName>] [-R] [-m] [-h]
```

例:

```
emrwal deleteWal -w examplews -p hbaseable -n examplewal
```

## deleteWorkspace

deleteWorkspace コマンドは、指定した Amazon EMR WAL ワークスペースを削除します。

使用:

```
emrwal deleteWorkspace [-e {endpoint}] [-r {Region}] -w {workspacename} [-h]
```

例:

```
emrwal deleteWorkspace -w examplews
```

## listTagsForResource

listTagsForResource コマンドは、指定した Amazon EMR WAL ワークスペースのすべてのキーと値のペアタグを一覧表示します。

使用:

```
emrwal listTagsForResource -arn {resource-arn} [-e {endpoint}] [-r {Region}] [-h]
```

例:

```
emrwal listTagsForResource -arn arn:aws:emrwal::1234567891234:workspace/examplews
```

## listWals

listWals コマンドは、指定したワークスペース内のすべての Amazon EMR WALs を一覧表示します。

使用:

```
emrwal listWals [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}] [-r {Region}] [-w {workspacename}] [-p <tablePrefix>] [-M {integer}] [-h]
```

例:

```
emrwal listWals -w examples
```

## listWorkspaces

listWorkspaces コマンドは、使用可能なすべての Amazon EMR WAL ワークスペースを一覧表示します。

使用:

```
emrwal listWorkspaces [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}] [-r {Region}] [-M {integer}] [-h]
```

例:

```
emrwal listWorkspaces
```

## tagResource

tagResource コマンドは、指定した Amazon EMR WAL ワークスペースに 1 つ以上のキーと値のペアタグを割り当てます。

使用:

```
emrwal tagResource -arn {resource-arn} -tags <tags> [-e {endpoint}] [-r {Region}] [-h]
```

例:

```
emrwal tagResource -arn arn:aws:emrwal::1234567891234:workspace/examples -  
tags tag_key=tag_value
```

## untagResource

untagResource コマンドは、指定した Amazon EMR WAL ワークスペースに 1 つ以上のキーと値のペアタグを割り当て解除します。

使用:

```
emrwal untagResource -arn {resource-arn} -tagKeys <tagKeys> [-e {endpoint}] [-  
r {Region}] [-h]
```

例:

```
emrwal untagResource -arn arn:aws:emrwal::1234567891234:workspace/examples -  
tagKeys tag_key
```

## HBase シェルを使用する

HBase クラスターを作成したら、次のステップとして HBase に接続してデータの読み書きを開始できます (データの書き込みはリードレプリカクラスターではサポートされません)。[HBase シェル](#)を使用すると、コマンドをテストできます。

HBase シェルを開くには

1. SSH を使用して HBase クラスターのメインサーバーに接続します。SSH を使用して Amazon EMR プライマリノードに接続する方法については、「[Amazon EMR 管理ガイド](#)」の「[SSH を使用してプライマリノードに接続する](#)」を参照してください。
2. `hbase shell` を実行します。HBase シェルが開き、次のようなプロンプトが表示されます。

```
hbase(main):001:0>
```

Base シェルコマンドを発行するには、プロンプトを使用します。シェルコマンドの説明や、シェルコマンドを呼び出す方法については、HBase プロンプトで「help」と入力し、Enter キーを押します。



## テーブルを作成する

次のコマンドを実行すると、「f1」という名前の 1 つの列ファミリーが含まれる、「t1」という名前のテーブルが作成されます。

```
hbase(main):001:0>create 't1', 'f1'
```

## 値を入力する

次のコマンドを実行すると、テーブル「t1」の行「r1」、および列「f1」に対して値「v1」が入力されます。

```
hbase(main):001:0>put 't1', 'r1', 'f1:col1', 'v1'
```

## 値を取得する

次のコマンドを実行すると、テーブル「t1」の行「r1」の値を取得します。

```
hbase(main):001:0>get 't1', 'r1'
```

## テーブルを削除する

次のコマンドはテーブル「t1」を削除します。

```
hbase(main):001:0>drop 'ns1:t1',false
```

boolean 値はテーブルをアーカイブするかどうかに対応しているので、保存する場合は true に設定できます。また、boolean を指定せずに drop 'ns1:t1' を実行してテーブルをアーカイブすることもできます。

## Hive で HBase テーブルにアクセスする

HBase と [Apache Hive](#) は、緊密に統合されているため、HBase に格納されたデータに対して並列処理の大量ワークロードを直接実行できます。HBase と共に Hive を使用するには、通常は同じクラスターでこれらを起動します。ただし Hive と HBase を別々のクラスターで起動することもできます。HBase と Hive を別々のクラスターに切り離して実行すると、各アプリケーションがクラスターのリソースをより効率的に使用できるため、パフォーマンスが向上する可能性があります。

次の手順は、Hive を使用してクラスターの HBase に接続する方法を説明します。

**Note**

Hive クラスターは 1 つの HBase クラスターにのみ接続できます。

Hive を HBase に接続するには

1. Hive と HBase がインストールされた別々のクラスターを作成するか、HBase と Hive の両方インストールされている 1 つのクラスターを作成します。
2. 別のクラスターを使用している場合は、HBase と Hive ポートがこれら 2 つのプライマリノード間で開かれるようにセキュリティグループを変更します。
3. SSH を使用して、Hive がインストールされたクラスターのプライマリノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してプライマリノードに接続する](#)」を参照してください。
4. 次のコマンドで Hive シェルを起動します。

```
hive
```

5. (オプション) HBase と Hive が同一のクラスターにある場合、これを行う必要はありません。Hive クラスターの HBase クライアントを、データが含まれる HBase クラスターに接続します。次の例では、*public-DNS-name* は HBase クラスターのプライマリノードのパブリック DNS 名に置き換えられます。例: `ec2-50-19-76-67.compute-1.amazonaws.com`。

```
set hbase.zookeeper.quorum=public-DNS-name;
```

6. 必要に応じて、HBase データに対して Hive クエリを実行するか、次の手順に進みます。

Hive から HBase データにアクセスするには

- Hive クラスターと HBase クラスター間の接続が確立されたら (前述の手順を参照)、Hive に外部テーブルを作成することで、HBase クラスターに格納されているデータにアクセスできます。

次の例では、プライマリノードの Hive プロンプトから実行すると、`h` という HBase テーブルに保存されているデータを参照する外部テーブルを作成します `inputTable`。その後、Hive ス

テートメントで `inputTable` を参照し、HBase クラスターに格納されているデータに対してクエリを実行したり、そのデータを変更したりできます。

```
set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com;

create external table inputTable (key string, value string)
  stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
  with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
  tblproperties ("hbase.table.name" = "t1");

select count(key) from inputTable ;
```

HBase と Hive を組み合わせたより高度なユースケースと例については、AWS ビッグデータブログ記事「[Amazon EMR で Apache HBase と Apache Hive を使用して NoSQL と超並列分析を組み合わせる](#)」を参照してください。

## HBase のスナップショットを使用する

HBase では、組み込みの[スナップショット](#)機能を使用して、テーブルの軽量バックアップが作成されます。EMR クラスターでは、EMRFS を使用してこれらのバックアップを Amazon S3 にエクスポートすることもできます。HBase シェルを使用して、プライマリノードにスナップショットを作成できます。このトピックでは、シェルでこれらのコマンドをインタラクティブに実行する方法、または AWS CLI または `command-runner.jar` を使用してステップを実行する方法について説明します AWS SDK for Java。HBase バックアップの他の種類については、HBase のドキュメントで「[HBase Backup](#)」の項を参照してください。

## テーブルを使用してスナップショットを作成する

```
hbase snapshot create -n snapshotName -t tableName
```

からの `command-runner.jar` の使用 AWS CLI :

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \  
--steps Name="HBase Shell Step",Jar="command-runner.jar",\  
Args=[ "hbase", "snapshot", "create", "-n", "snapshotName", "-t", "tableName"]
```

## AWS SDK for Java

```
HadoopJarStepConfig hbaseSnapshotConf = new HadoopJarStepConfig()  
.withJar("command-runner.jar")  
.withArgs("hbase","snapshot","create","-n","snapshotName","-t","tableName");
```

### Note

スナップショット名が一意でない場合、作成処理はリターンコード -1 または 255 で失敗になりますが、問題点を記述したエラーメッセージが表示されない可能性があります。同じスナップショット名を使用するには、いったん削除してから再作成してください。

## スナップショットを削除する

```
hbase shell  
>> delete_snapshot 'snapshotName'
```

## スナップショット情報を表示する

```
hbase snapshot info -snapshot snapshotName
```

## Amazon S3 にスナップショットをエクスポートする

### Important

スナップショットをエクスポートするときに `-mappers` 値を指定しないと、HBase は任意の計算を使用してマッパーの数を判断します。この値は、テーブルサイズによっては非常に大きくなる場合があります、エクスポート時のジョブの実行に悪影響を及ぼします。したがって、`-mappers` パラメータ、`-bandwidth` パラメータ (帯域幅の消費をメガバイト/秒で指定)、または両方を指定してエクスポートオペレーションで使用されるクラスターリソースを制限することをお勧めします。または、使用量が低下する期間にスナップショットのエクスポートオペレーションを実行できます。

```
hbase snapshot export -snapshot snapshotName \  
-copy-to s3://bucketName/folder -mappers 2
```

`command-runner.jar` から 使用する AWS CLI :

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
Args=[ "hbase", "snapshot", "export", "-snapshot", "snapshotName", "-copy-
to", "s3://bucketName/folder", "-mappers", "2", "-bandwidth", "50"]
```

AWS SDK for Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("hbase", "snapshot", "export",
        "-snapshot", "snapshotName", "-copy-to",
        "s3://bucketName/folder",
        "-mappers", "2", "-bandwidth", "50");
```

## Amazon S3 からスナップショットをインポートする

操作はインポートですが、ここで使用する HBase は export です。

```
sudo -u hbase hbase snapshot export \
-D hbase.rootdir=s3://bucketName/folder \
-snapshot snapshotName \
-copy-to hdfs://masterPublicDNSName:8020/user/hbase \
-mappers 2
```

command-runner.jar から を使用する AWS CLI :

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar", \
Args=["sudo", "-u", "hbase", "hbase snapshot export", "-snapshot", "snapshotName", \
"-D", "hbase.rootdir=s3://bucketName/folder", \
"-copy-to", "hdfs://masterPublicDNSName:8020/user/hbase", "-mappers", "2", "-chmod", "700"]
```

AWS SDK for Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("sudo", "-u", "hbase", "hbase", "snapshot", "export", "-D", "hbase.rootdir=s3://
path/to/snapshot",
        "-snapshot", "snapshotName", "-copy-to",
        "hdfs://masterPublicDNSName:8020/user/hbase",
        "-mappers", "2", "-chuser", "hbase");
```

## HBase シェル内のスナップショットからテーブルを復元する

```
hbase shell
>> disable tableName
>> restore_snapshot snapshotName
>> enable tableName
```

HBase では現在、HBase シェルで見つかるスナップショットコマンドがすべてサポートされていません。たとえば、スナップショットを復元する HBase コマンドラインオプションはないため、シェルで復元する必要があります。つまり、`command-runner.jar` で Bash コマンドを実行する必要があります。

### Note

ここで使用されているコマンドは `echo` であるため、Amazon EMR で実行されたコマンドが終了コード 0 を返しても、シェルコマンドは失敗する可能性があります。ステップとしてシェルコマンドを実行する場合は、ステップログを確認してください。

```
echo 'disable tableName; \
restore_snapshot snapshotName; \
enable tableName' | hbase shell
```

AWS CLIを使用するステップを次に示します。まず、以下の内容 `snapshot.json` ファイルを作成します。

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo '$'disable \"tableName\"; restore_snapshot \"snapshotName\"; enable \"tableName\" | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file:///./snapshot.json
```

## AWS SDK for Java:

```
HadoopJarStepConfig hbaseRestoreSnapshotConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("bash", "-c", "echo '$disable \"tableName\"; restore_snapshot \"snapshotName\"; enable \"snapshotName\"' | hbase shell");
```

## HBase を設定する

デフォルトの HBase 設定はほとんどのアプリケーションで動作しますが、HBase 設定を変更することもできます。そのためには、HBase 設定分類のプロパティを使用します。詳細については、「[アプリケーションの設定](#)」を参照してください。

次の例では、Amazon S3 に格納されている設定ファイル `myConfig.json` に基づいて、代替の HBase ルートディレクトリを持つクラスターを作成します。

### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --release-label emr-7.1.0 --applications Name=HBase \
--instance-type m5.xlarge --instance-count 3 --configurations https://s3.amazonaws.com/
mybucket/myfolder/myConfig.json
```

`myConfig.json` ファイルは、次の例に示すように、`hbase.rootdir` 設定分類の `hbase-site` プロパティを指定します。`ip-XXX-XX-XX-XXX.ec2.internal` をクラスターのプライマリノードの内部 DNS ホスト名に置き換えます。

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "hbase.rootdir": "hdfs://ip-XXX-XX-XX-XXX.ec2.internal:8020/user/
myCustomHBaseDir"
    }
  }
]
```

]

**Note**

Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## YARN でメモリ割り当てを変更する

HBase は YARN アプリケーションとして実行されていないため、YARN および YARN アプリケーションに割り当てられたメモリを再計算する必要があります。その結果、HBase がインストールされている場合は YARN で全体的に使用可能なメモリ量が少なくなります。YARN アプリケーションと HBase を同じクラスターに配置することを計画する場合は、この点を考慮する必要があります。64 GB 未満のメモリを持つインスタンスタイプは、で使用可能なメモリの半分を持ち NodeManager、HBase RegionServer に割り当てられます。メモリが 64 GB を超えるインスタンスタイプの場合、HBase RegionServer メモリは 32 GB に制限されます。原則として、YARN 設定メモリは MapReduce リデューサータスクメモリの倍数です。

「[タスクの構成設定のデフォルト値](#)」のテーブルには、HBase に必要なメモリに基づいた YARN 設定の変更が示されています。

## HBase のポート番号

HBase 用に選択されている一部のポート番号は、デフォルト値と異なります。Amazon EMR の HBase 用インスタンスおよびポートを次に示します。

### HBase のポート

インターフェイス	[ポート]	[プロトコル]
HMaster	16000	TCP
HMaster UI	16010	HTTP
RegionServer	16020	TCP



インターフェイス	[ポート]	[プロトコル]
RegionServer 情報	16030	HTTP
REST サーバー	8070	HTTP
REST UI	8085	HTTP
Thrift サーバー	9090	TCP
Thrift サーバー UI	9095	HTTP

### Important

Amazon EMR リリースバージョン 4.6.0 以降では、kms-http-port は 9700、kms-admin-port は 9701 です。

## 最適化のための HBase サイト設定

HBase サイト設定の一部またはすべてを設定することで、アプリケーションのワークロード用に HBase クラスターを最適化できます。調査の出発点として次の設定を使用することをお勧めします。

### zookeeper.session.timeout

デフォルトのタイムアウトは 40 秒 (40000 ミリ秒) です。リージョンサーバーがクラッシュしたとき、そのリージョンサーバーが存在しないことをマスターサーバーが認識し、回復処理を開始するのに要する時間です。マスターサーバーの回復処理を高速化するには、この値を小さくして時間を短くします。次の例では、30 秒、つまり 30000 ミリ秒を使用しています。

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "zookeeper.session.timeout": "30000"
    }
  }
]
```

## hbase.regionserver.handler.count

リクエストをテーブルに提供するためにリージョンサーバーが開いたままにするスレッド数を定義します。同時実行クライアント数が多い大きな書き込みバッファを使用するときに、ユーザーによるリージョンサーバーの強制終了を防ぐため、デフォルトでは 10 という低い値が設定されています。大まかな方法として、リクエストあたりのペイロードが MB 範囲に近いときは ( big puts、ラージキャッシュを使用した scans ) 小さな値を、ペイロードが小さいときは ( gets、small puts、ICV、deletes )、大きな値を指定します。次の例では、開いているスレッドの数を 30 まで増やしています。

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "hbase.regionserver.handler.count": "30"
    }
  }
]
```

## hbase.hregion.max.filesize

このパラメータは、個別のリージョンのサイズ ( バイト単位 ) を管理します。デフォルトでは、1073741824 に設定されます。大量のデータを HBase クラスターに書き込んでおり、頻繁に分割が発生する場合は、このサイズを増やすことで個別のリージョンを大きくします。分割は少なくなりますが、サーバー間におけるリージョンの負荷分散にかかる時間が長くさなります。

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "hbase.hregion.max.filesize": "1073741824"
    }
  }
]
```

## hbase.hregion.memstore.flush.size

このパラメータは、ディスクにフラッシュする前の memstore の最大サイズ ( バイト単位 ) を管理します。デフォルトでは、134217728 です。ワークロードが書き込みオペレーションの短期バースト

で構成されている場合は、この制限を増やして、バースト中もすべての書き込みがメモリに収まるようにし、後でディスクにフラッシュされるようにします。これにより、バースト中のパフォーマンスが向上することがあります。

```
[
  {
    "Classification": "hbase-site",
    "Properties": {
      "hbase.hregion.memstore.flush.size": "134217728"
    }
  }
]
```

## HBase ユーザーインターフェイスを表示する

### Note

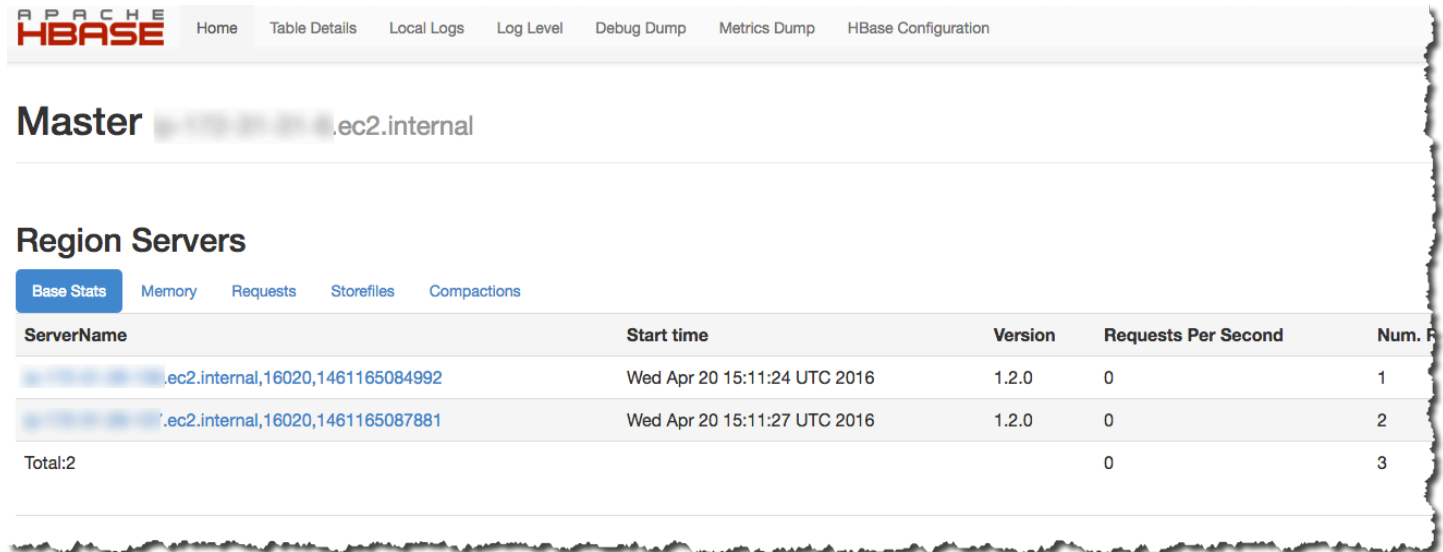
HBase ユーザーインターフェイスは、デフォルトで安全ではない HTTP 接続を使用します。セキュア HTTP (HTTPS) を有効にするには、[HBase 構成](#)で hbase-site 分類の hbase.ssl.enabled プロパティを true に設定します。HBase ウェブ UI でセキュア HTTP (HTTPS) を使用方法の詳細については、「[Apache HBase リファレンスガイド](#)」を参照してください。

HBase にはウェブベースのユーザーインターフェイスが用意されており、これを使用して HBase クラスタをモニタリングできます。Amazon EMR で HBase を実行すると、ウェブインターフェイスはプライマリノードで実行され、SSH トンネルの作成とも呼ばれるポート転送を使用して表示できます。

HBase ユーザーインターフェイスを表示するには

1. SSH を使用してプライマリノードにトンネルし、安全な接続を作成します。詳細については、「Amazon EMR 管理ガイド」の「[オプション 2、パート 1: 動的ポート転送を使用してプライマリノードへの SSH トンネルを設定する](#)」を参照してください。
2. Firefox 用 FoxyProxy プラグインなどのプロキシツールを使用してウェブブラウザをインストールし、AWS ドメインの SOCKS プロキシを作成します。詳細については、「Amazon EMR 管理ガイド」の「[オプション 2、パート 2: プライマリノードでホストされているウェブサイトを表示するようにプロキシ設定を構成する](#)」を参照してください。

3. プロキシを設定し、SSH 接続を開いた状態で、`http://master-public-dns-name:16010/master-status` でブラウザウィンドウを開くことで HBase UI を表示できます。ここで、**master-public-dns-name** はクラスターのプライマリノードのパブリック DNS アドレスです。



APACHE HBASE Home Table Details Local Logs Log Level Debug Dump Metrics Dump HBase Configuration

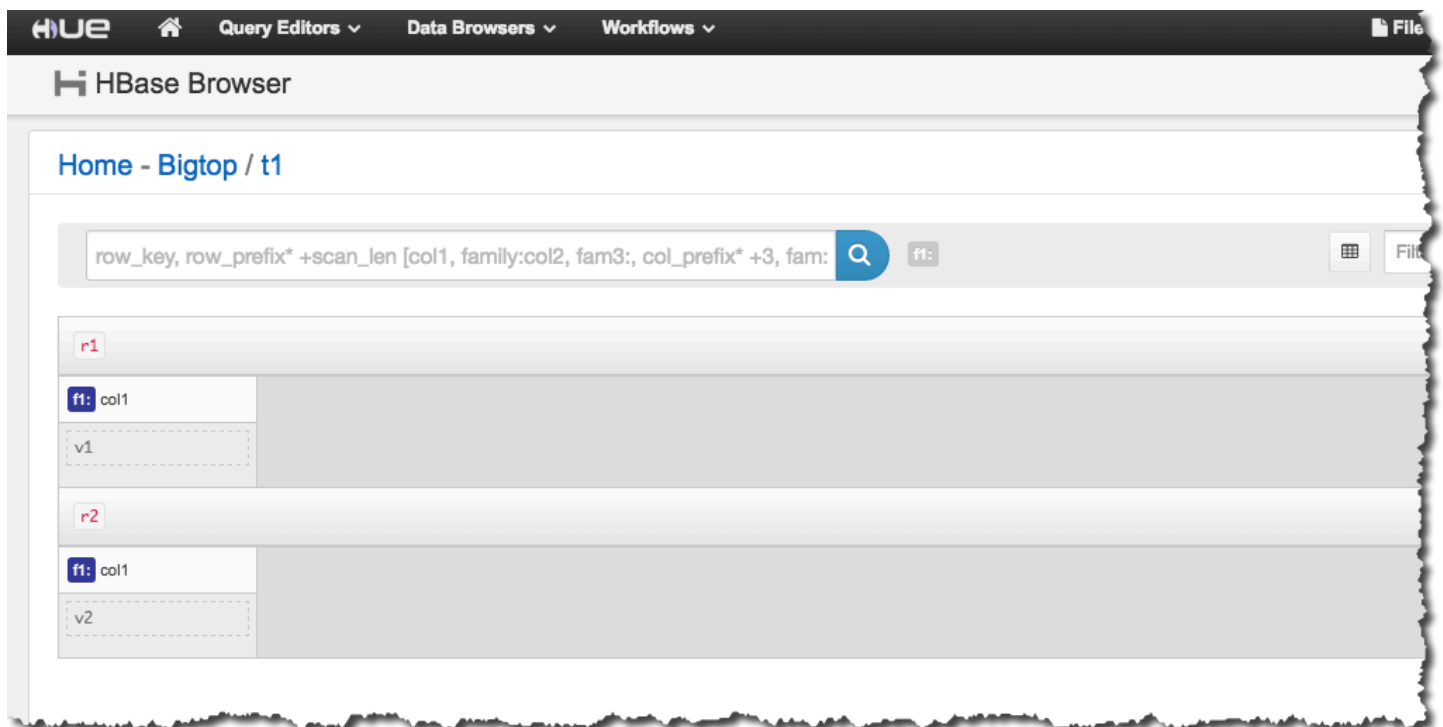
## Master **ec2.internal**

### Region Servers

Base Stats Memory Requests Storefiles Compactions

ServerName	Start time	Version	Requests Per Second	Num. P
<b>ec2.internal,16020,1461165084992</b>	Wed Apr 20 15:11:24 UTC 2016	1.2.0	0	1
<b>ec2.internal,16020,1461165087881</b>	Wed Apr 20 15:11:27 UTC 2016	1.2.0	0	2
Total:2			0	3

HBase は Hue で表示することもできます。たとえば、次の図では、「t1」で作成したテーブル [HBase シェルを使用する](#) が示されています。



HUE Query Editors Data Browsers Workflows File

## HBase Browser

Home - Bigtop / t1

row\_key, row\_prefix\* +scan\_len [col1, family:col2, fam3:, col\_prefix\* +3, fam: Q ft: Filter

r1	f1: col1	v1
r2	f1: col1	v2

Hue の詳細については、「[Hue](#)」を参照してください。

## HBase ログファイルを表示する

HBase では、そのオペレーションの一環として、設定、デーモン、アクション、および例外に関する詳しい情報をログファイルに書き込みます。これらのログファイルは、HBase で問題をデバッグするだけでなく、パフォーマンスを追跡するときにも役に立ちます。

ログファイルを Amazon S3 に保持するようクラスターを設定した場合、ログが Amazon S3 に書き込まれるのは 5 分ごとであり、最新のログファイルが利用できるようになるまでに若干の遅れが生じる可能性があるという点に注意する必要があります。

プライマリノードで HBase ログを表示するには

- SSH を使用してプライマリノードに接続し、`/var/log/hbase` ディレクトリに移動することで、現在の HBase ログを表示できます。これらのログは、クラスターの起動時に Amazon S3 へのログを有効にしない限り、クラスター終了後に使用できなくなります。

Amazon S3 で HBase ログを表示するには

- Amazon S3 で HBase ログおよび他のクラスターログにアクセスするには、また、クラスターの終了後もこうしたログを利用できるようにするには、クラスターの作成時にこれらのログを受け取るよう Amazon S3 バケットを指定する必要があります。これは、`--log-uri` オプションを使用して行います。クラスターのログ記録を有効にする方法の詳細については、「Amazon EMR 管理ガイド」の「[ログ記録およびデバッグの設定 \(オプション\)](#)」を参照してください。

## Ganglia で HBase を監視する

Ganglia オープンソースプロジェクトは拡張可能な分散システムで、パフォーマンスへの影響を最小限に抑えながら、クラスターやグリッドをモニタリングできるように設計されています。クラスターで Ganglia を有効にすると、レポートを生成し、クラスター全体のパフォーマンスを表示するだけでなく、個別のノードインスタンスのパフォーマンスを調べることができます。Ganglia オープンソースプロジェクトの詳細については、<http://ganglia.info/> を参照してください。Amazon EMR クラスターでの Ganglia の使用の詳細については、「[Ganglia](#)」を参照してください。

Ganglia を設定してクラスターを起動すると、プライマリノードで実行されているグラフィカルインターフェイスを使用して Ganglia グラフとレポートにアクセスできます。

Ganglia はログファイルを `/mnt/var/lib/ganglia/rrds/` ディレクトリのプライマリノードに保存します。Amazon EMR の以前のリリースバージョンは `/var/log/ganglia/rrds/` ディレクトリにログファイルを保存する場合があります。

を使用して Ganglia と HBase のクラスターを設定するには AWS CLI

- 以下のような `create-cluster` コマンドを使用します。

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.1.0 \  
--applications Name=HBase Name=Ganglia --use-default-roles \  
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \  
--instance-count 3
```

#### Note

デフォルトの Amazon EMR サービスロールと Amazon EC2 インスタンスプロファイルが存在しない場合、エラーが発生します。aws emr create-default-roles コマンドを使用して作成してから、もう一度お試しください。

詳細については、「[AWS CLIの Amazon EMR コマンド](#)」を参照してください。

Ganglia のウェブインターフェイスで HBase メトリクスを表示するには

1. SSH を使用してプライマリノードにトンネルし、安全な接続を作成します。詳細については、「Amazon EMR 管理ガイド」の「[オプション 2、パート 1: 動的ポート転送を使用してプライマリノードへの SSH トンネルを設定する](#)」を参照してください。
2. Firefox 用 FoxyProxy プラグインなどのプロキシツールを使用してウェブブラウザをインストールし、AWS ドメインの SOCKS プロキシを作成します。詳細については、「Amazon EMR 管理ガイド」の「[オプション 2、パート 2: プライマリノードでホストされているウェブサイトを表示するようにプロキシ設定を構成する](#)」を参照してください。
3. プロキシを設定し、SSH 接続を開いた状態で、`http://master-public-dns-name/ganglia/` でブラウザウィンドウを開くことで Ganglia メトリクスを表示できます。ここで、`master-public-dns-name` は HBase クラスター内のマスターサーバーのパブリック DNS アドレスです。

## プライマリノードで Ganglia ログファイルを表示するには

- クラスターがまだ実行中の場合は、SSH を使用してプライマリノードに接続し、`/mnt/var/lib/ganglia/rrds/` ディレクトリに移動することでログファイルにアクセスできます。EMR 3.x の場合、`/var/log/ganglia/rrds` ディレクトリに移動します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してプライマリノードに接続する](#)」を参照してください。

## Amazon S3 で Ganglia ログファイルを表示するには

- クラスターのログを有効にしても、Ganglia ログファイルは自動的に Amazon S3 に書き込まれません。Amazon S3 で Ganglia ログファイルを表示するには、`/mnt/var/lib/ganglia/rrds/` から S3 バケットにログを手動でプッシュする必要があります。

## HBase の以前のバージョンからの移行

HBase の以前のバージョンからデータを移行するには、Apache HBase リファレンスガイドの「[Upgrading](#)」および「[HBase version number and compatibility](#)」を参照してください。HBase の 1.0 バージョンより以前のバージョンからアップグレードする場合は、要件に特にご注意ください。

## HBase リリース履歴

次の表は、Amazon EMR の各リリースに含まれている HBase のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### HBase バージョン情報

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.36.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
		de, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-7.1.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-7.0.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされる コンポーネント
emr-6.15.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-ma-pred, hadoop-yarn-nodema-nager, hadoop-yarn-resour-cemanager, hadoop-yarn-timeline-server, hbase-hma-ster, hbase-client, hbase-reg-ion-server, hbase-rest-server, hbase-thrift-server, hbase-ope-rator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.14.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema-nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.13.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-ma-pred, hadoop-yarn-nodema-nager, hadoop-yarn-resour-cemanager, hadoop-yarn-timeline-server, hbase-hma-ster, hbase-client, hbase-reg-ion-server, hbase-rest-server, hbase-thrift-server, hbase-ope-rator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.12.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema-nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされる コンポーネント
emr-6.11.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema-nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.11.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema-nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.10.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema-nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.10.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano-de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema-nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.9.1	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.9.0	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.8.1	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.8.0	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.7.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.36.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.6.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.35.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.5.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.4.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.3.1	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.3.0	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.2.1	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.2.0	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.1.1	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.1.0	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.0.1	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-6.0.0	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.34.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.33.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.33.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.32.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.32.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.31.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.31.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.30.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.30.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.30.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.29.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.28.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.28.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.27.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.27.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.26.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.25.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.24.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.24.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.23.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.23.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.22.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.2	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.21.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.20.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.20.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.17.2	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.17.1	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.17.0	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.16.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.16.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.15.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.15.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.2	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.14.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.13.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.13.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.12.3	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.2	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.12.1	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.0	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.11.4	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.11.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.11.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.10.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.10.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.9.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.9.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.8.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.8.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.7.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.7.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.6.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.6.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.4	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.5.3	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.2	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.5.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.4.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.4.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.3.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.3.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.3.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.3	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.2.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.2.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.1.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.1.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.0.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-5.0.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.6	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.9.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.9.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.9.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.8.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.8.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.8.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.4	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.7.3	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.2	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.7.1	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.0	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	HBase バージョン	HBase でインストールされるコンポーネント
emr-4.6.1	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.6.0	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

# Apache HCatalog

HCatalog は、Pig、Spark SQL、および/またはカスタム MapReduce アプリケーション内の Hive メタストアテーブルにアクセスできるようにするツールです。HCatalog には、テーブルの作成やその他のオペレーションに使用できる REST インターフェイスおよびコマンドラインクライアントがあります。その後、HCatalog ライブラリを使用してテーブルにアクセスするアプリケーションを記述します。詳細については、「[Using HCatalog](#)」を参照してください。HCatalog は、Amazon EMR リリースバージョン 4.4.0 以降に含まれています。

Amazon EMR リリースバージョン 5.8.0 以降の HCatalog では、Hive のメタストアとして AWS Glue Data Catalog の使用がサポートされています。詳細については、「[Hive のメタストアとしての AWS Glue データカタログの使用](#)」を参照してください。

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている HCatalog のバージョンと、Amazon EMR で HCatalog と共にインストールされるコンポーネントを示しています。

このリリースで HCatalog と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の HCatalog バージョン情報

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.15.0	HCatalog 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている HCatalog のバージョンと、Amazon EMR で HCatalog と共にインストールされるコンポーネントを示しています。

このリリースで HCatalog と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

#### emr-5.36.2 の HCatalog バージョン情報

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.36.2	HCatalog 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

#### トピック

- [HCatalog を使用したクラスターの作成](#)
- [HCatalog の使用](#)
- [例: HCatalog テーブルを作成して Pig を使用して書き込む](#)
- [HCatalog リリース履歴](#)

## HCatalog を使用したクラスターの作成

HCatalog は Hive プロジェクトに含まれていますが、お客様のアプリケーションとしてインストールする必要があります。



コンソールを使用し、HCatalog がインストールされた状態でクラスターを起動するには

次の手順では、HCatalog がインストールされたクラスターを作成します。詳細オプションなど、コンソールを使用したクラスター作成の詳細については、「[Amazon EMR 管理ガイド](#)」の「クラスターの計画と構成」を参照してください。

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Create cluster (クラスターの作成)] を選択して [Quick Create (クイック作成)] を使用します。
3. [Software Configuration (ソフトウェア設定)] フィールドで、Amazon Release Version emr-4.4.0 以降を選択します。
4. [Select Applications (アプリケーションの選択)] フィールドで、[All Applications (すべてのアプリケーション)] または [HCatalog] を選択します。
5. 必要に応じて他のオプションを選択し、[Create cluster (クラスターの作成)] を選択します。

を使用して HCatalog でクラスターを起動するには AWS CLI

- 次のコマンドを使用してクラスターを作成します。

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Cluster with Hcat" --release-label emr-7.1.0 \  
--applications Name=HCatalog --ec2-attributes KeyName=myKey \  
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

## HCatalog の使用

Hive メタストアを使用するさまざまなアプリケーション内で HCatalog を使用できます。このセクションの例は、テーブルを作成し、それを Pig および Spark SQL のコンテキストで使用する方法を示しています。

## HCatalog HStorer の使用時に直接書き込みを無効にする

アプリケーションが [HCatStorer](#) を使用して Amazon S3 に保存されている HCatalog テーブルに書き込む場合は、Amazon EMR の直接書き込み機能を無効にします。Amazon S3 例えば、Pig の STORE コマンドを使用するときや、HCatalog テーブルを Amazon S3 に書き込む Sqoop ジョブを実行するとき、直接書き込みを無効にします。直接書き込み機能を無効にするには、`mapred.output.direct.NativeS3FileSystem` および `mapred.output.direct.EmrFileSystem` を `false` に設定します。次の例は、Java を使用してこれらの構成を設定する方法を示しています。

```
Configuration conf = new Configuration();
conf.set("mapred.output.direct.NativeS3FileSystem", "false");
conf.set("mapred.output.direct.EmrFileSystem", "false");
```

## HCat CLI を使用してテーブルを作成しそのデータを Pig で使用する

クラスターで次のスクリプト (`impressions.q`) を作成します。

```
CREATE EXTERNAL TABLE impressions (
  requestBeginTime string, adId string, impressionId string, referrer string,
  userAgent string, userCookie string, ip string
)
PARTITIONED BY (dt string)
ROW FORMAT
  serde 'org.apache.hive.hcatalog.data.JsonSerDe'
  with serdeproperties ( 'paths'='requestBeginTime, adId, impressionId, referrer,
userAgent, userCookie, ip' )
LOCATION 's3://[your region].elasticmapreduce/samples/hive-ads/tables/impressions/';
ALTER TABLE impressions ADD PARTITION (dt='2009-04-13-08-05');
```

HCat CLI を使用してスクリプトを実行します。

```
% hcat -f impressions.q
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-
log4j.properties
OK
Time taken: 4.001 seconds
OK
Time taken: 0.519 seconds
```

Grunt シェルを開き、`impressions` にあるデータにアクセスします。

```
% pig -useHCatalog -e "A = LOAD 'impressions' USING
  org.apache.hive.hcatalog.pig.HCatLoader();
B = LIMIT A 5;
dump B;"
<snip>
(1239610346000,m9nwdo67Nx6q2kI25qt50n7peICfUM,omkxkaRpNhGPDucAiBErSh1cs0MThC,cartoonnetwork.com
  (compatible; MSIE 7.0; Windows NT 6.0; FunWebProducts; GTB6; SLCC1; .NET CLR
  2.0.50727; Media Center PC
  5.0; .NET,wcVWWTascoPbGt6bdqDbuWTPPHgOPs,69.191.224.234,2009-04-13-08-05)
(1239611000000,NjriQjd0DgWBKnkGJUP6GNTbDeK4An,AwtXPkfaWG0aNeL900sFU8Hcj6eLHT,cartoonnetwork.com
  (compatible; MSIE 7.0; Windows NT 5.1; GTB6; .NET CLR
  1.1.4322),0aMU1F2gE4CtADVHAbKjjRRks5kIgg,57.34.133.110,2009-04-13-08-05)
(1239610462000,Irpv3oiu0I5QNQiWSSITshrLdo9cM1,i1LDq44LRSJF0hbmhB8Gk7k9gMWtBq,cartoonnetwork.com
  (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
  InfoPath.1),Qsb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239611007000,q2Aawfnpe0JAvhInaIp0VGx9KTS0oPO,s3HvTf1PB8JIE0IuM6h0EebWwP0tJV,cartoonnetwork.com
  (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
  InfoPath.1),Qsb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239610398000,c362vpAB0soPKGHRs43cj6TRwNe0Gn,jeas5nXbQInGAgFB8jlkhnprN6cMw7,cartoonnetwork.com
  (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0; GTB6; .NET CLR
  1.1.4322),k96n5PnUmwHKfiUI0TFP0TNMFADgh9,51.131.29.87,2009-04-13-08-05)
7120 [main] INFO  org.apache.pig.Main - Pig script completed in 7 seconds and 199
  milliseconds (7199 ms)
16/03/08 23:17:10 INFO pig.Main: Pig script completed in 7 seconds and 199 milliseconds
  (7199 ms)
```

## Spark SQL を使用してテーブルにアクセスする

この例では、最初の例で作成したテーブル DataFrame から Spark を作成し、最初の 20 行を示します。

```
% spark-shell --jars /usr/lib/hive-hcatalog/share/hcatalog/hive-hcatalog-core-1.0.0-
  amzn-3.jar
<snip>
scala> val hiveContext = new org.apache.spark.sql.hive.HiveContext(sc);
scala> val df = hiveContext.sql("SELECT * FROM impressions")
scala> df.show()
<snip>
16/03/09 17:18:46 INFO DAGScheduler: ResultStage 0 (show at <console>:32) finished in
  10.702 s
16/03/09 17:18:46 INFO DAGScheduler: Job 0 finished: show at <console>:32, took
  10.839905 s
```

```

+-----+-----+-----+-----+
+-----+-----+-----+-----+
|requestbegintime|          adid|          impressionid|          referrer|
|  useragent|          usercookie|          ip|          dt|
+-----+-----+-----+-----+
+-----+-----+-----+-----+
|  1239610346000|m9nwd067Nx6q2kI25...|omkxkaRpNhGPDucAi...|cartoonnetwork.com|
Mozilla/4.0 (comp...|wcVWWTascoPbGt6bd...|69.191.224.234|2009-04-13-08-05|
|  1239611000000|NjriQjd0DgWBKnkGJ...|AWtXPKfawG0aNeL90...|cartoonnetwork.com|
Mozilla/4.0 (comp...|0aMU1F2gE4CtADVHA...| 57.34.133.110|2009-04-13-08-05|
|  1239610462000|Irpv3oiu0I5QNQiW...|i1LDq44LRSJF0hbmh...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|  1239611007000|q2Awfnp0JAvhInaI...|s3HvTf1PB8JIE0IuM...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|  1239610398000|c362vpAB0soPKGHR...|jeas5nXbQInGAgFB8...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...| 51.131.29.87|2009-04-13-08-05|
|  1239610600000|cjbTpruoaiEtqLuMX...|XwlohBSs8Ipxs1bRa...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...| 51.131.29.87|2009-04-13-08-05|
|  1239610804000|Ms3eJHNAEItpxvimd...|4SIj4pGmgVL1625BD...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...| 51.131.29.87|2009-04-13-08-05|
|  1239610872000|h5bccHX6wJReDi1jL...|EFAWiiBdVfnxwAMWP...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...| 51.131.29.87|2009-04-13-08-05|
|  1239610365000|874NBpGmxNFfxEPKM...|xSvE4XtGbdTXPF2Lb...|cartoonnetwork.com|
Mozilla/5.0 (Maci...|eWDEVVUphlnRa273j...| 22.91.173.232|2009-04-13-08-05|
|  1239610348000|X8gISpUTSqh1A5reS...|TrFblGT99AgE75vuj...|          corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...| 55.35.44.79|2009-04-13-08-05|
|  1239610743000|kbKreLWB6QVueFrDm...|kVnxx9Ie2i30LTxFj...|          corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...| 55.35.44.79|2009-04-13-08-05|
|  1239610812000|9lX0SRpEi3bmEeTCu...|1B2sff99AEIwSuLVV...|          corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...| 55.35.44.79|2009-04-13-08-05|
|  1239610876000|lijjmCf2kuxfBTnjL...|AjvufgUtakUFcsIM9...|          corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...| 55.35.44.79|2009-04-13-08-05|
|  1239610941000|t8t8trgjNRPilmxuD...|agu2u2TCdqWP08rAA...|          corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...| 55.35.44.79|2009-04-13-08-05|
|  1239610490000|OGRLPVNGxiGgrCmWL...|mJg2raBUpPrC80lUm...|          corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...| 71.124.66.3|2009-04-13-08-05|
|  1239610556000|OnJID12x0RXKPUgrD...|P7Pm2mPdW6w08KA3R...|          corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...| 71.124.66.3|2009-04-13-08-05|
|  1239610373000|WflsvKIg0qfIE5KwR...|TJHd1VBspNcua0XPn...|          corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|  1239610768000|4MJR0XxiVCU1ueXKV...|10hGwmbvKf8ajoU8a...|          corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|  1239610832000|gWIrpDiN57i3sHatv...|RNL4C7xPi3tdar2Uc...|          corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|

```

```
| 1239610789000|pTne9k62kJ14QViXI...|RVxJVIQousjxUVI3r...| pixnet.net|
Mozilla/5.0 (Maci...|1bG0KiBD2xmui90kF...| 33.176.101.80|2009-04-13-08-05|
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
only showing top 20 rows

scala>
```

## 例: HCatalog テーブルを作成して Pig を使用して書き込む

HCatalog テーブルを作成し、Apache Pig を使用して、Amazon S3 のデータソース CatStorer を使用して H でそのテーブルに書き込むことができます。Amazon S3 HCatalog では、直接書き込みを無効にする必要があります。無効にしないと操作が警告なしに失敗します。および `mapred.output.direct.NativeS3FileSystem` 構成の両方を、`mapred.output.direct.EmrFileSystem` 分類を使用することで、`false` または Grunt シェル内から `mapred-site` 手動で設定します。次の例では、HCat CLI を使用して作成されたテーブルと、Amazon S3 内のサンプルデータファイルからテーブルに値を設定するために Grunt シェルで実行されるコマンドを示しています。

この例を実行するには、[SSH を使用してマスターノードに接続します](#)。

次のコンテンツで HCatalog スクリプトファイル `wikicount.q` を作成します。これにより、`wikicount` という名前の HCatalog テーブルが作成されます。

```
CREATE EXTERNAL TABLE IF NOT EXISTS wikicount(
  col1 string,
  col2 bigint
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\001'
STORED AS ORC
LOCATION 's3://MyBucket/hcat/wikicount';
```

HCat CLI コマンドを使用してファイルからスクリプトを実行します。

```
hcat -f wikicount.q
```

次に、`-useHCatalog` オプションで Grunt シェルを起動し、直接書き込みを無効にするように構成を設定し、S3 の場所からデータを読み込み、その結果を `wikicount` テーブルに書き込みます。

```

pig -useHCatalog
SET mapred.output.direct.NativeS3FileSystem false;
SET mapred.output.direct.EmrFileSystem false;
A = LOAD 's3://support.elasticmapreduce/training/datasets/wikistats_tiny/' USING
  PigStorage(' ') AS (Site:chararray, page:chararray, views:int, total_bytes:long);
B = GROUP A BY Site;
C = FOREACH B GENERATE group as col1, COUNT(A) as col2;
STORE C INTO 'wikicount' USING org.apache.hive.hcatalog.pig.HCatStorer();

```

## HCatalog リリース履歴

次の表は、Amazon EMR の各リリースに含まれている HCatalog のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### HCatalog のバージョン情報

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.36.2	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
		datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.36.1	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.36.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.35.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server



Amazon EMR リリースラベル	HCatalog バージョン	HCatalog でインストールされるコンポーネント
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

# Apache Hive

Hive はオープンソースのデータウェアハウス兼分析パッケージで、Hadoop クラスター上で動作します。Hive スクリプトは、Hive QL (クエリ言語) という SQL に似た言語を使用しています。この言語はプログラミングモデルを抽象化しており、データウェアハウスとの一般的なやりとりをサポートします。Hive を使用すると、有向非巡回グラフ (DAGs などの下位レベルのコンピュータ言語での MapReduce プログラムに基づいて Tez ジョブを記述する複雑さを回避できます。

Hive は、シリアライズ形式を取り込むことで、SQL の枠組みを拡張しています。また、データ自体は操作せずにデータと一致するテーブルスキーマを作成することで、クエリ処理をカスタマイズすることもできます。SQL は日付、数値、文字列などのプリミティブなデータ型のみサポートしていますが、Hive テーブルの値は JSON オブジェクト、ユーザー定義のデータ型、Java で記述した関数などの構造化された要素です。

Hive の詳細については、<http://hive.apache.org/> を参照してください。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Hive のバージョンと、Amazon EMR で Hive と共にインストールされるコンポーネントを示しています。

このリリースで Hive と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Hive バージョン情報

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-7.1.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client,

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
		hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Hive のバージョンと、Amazon EMR で Hive と共にインストールされるコンポーネントを示しています。

このリリースで Hive と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Hive バージョン情報

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.15.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Hive のバージョンと、Amazon EMR で Hive と共にインストールされるコンポーネントを示しています。

このリリースで Hive と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

#### emr-5.36.2 の Hive バージョン情報

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.36.2	Hive 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 5.18.0 以降では、Amazon EMR アーティファクトリポジトリを使用し、特定の Amazon EMR リリースで利用可能なライブラリおよび依存関係の正確なバージョンに対してジョブコードをビルドできます。詳細については、[「Amazon EMR アーティファクトリポジトリを使用して依存関係を確認する」](#)を参照してください。

#### トピック

- [Amazon EMR 上の Hive の相違点と考慮事項](#)
- [Hive の外部メタストアの設定](#)
- [Hive JDBC ドライバーの使用](#)
- [Hive のパフォーマンスの向上](#)
- [Hive Live Long and Process \(LLAP\) の使用](#)

- [Hive での暗号化](#)
- [Hive リリース履歴](#)

## Amazon EMR 上の Hive の相違点と考慮事項

### Amazon EMR 上の Apache Hive と Apache Hive の相違点

このセクションでは、Amazon EMR 上の Hive と <http://svn.apache.org/viewvc/hive/branches/> で入手できるデフォルトバージョンの Hive の相違点について説明します。

#### Hive 認可

Amazon EMR は、EMRFS と Amazon S3 ではなく HDFS の [Hive 認可](#) をサポートします。Amazon EMR クラスタは、デフォルトで認可が無効な状態で実行されます。

#### Amazon S3 での Hive ファイルマージ動作

Apache Hive は `hive.merge.mapfiles` が `true` の場合はマップのみジョブの最後に小さいファイルをマージし、マージはジョブの平均出力サイズが `hive.merge.smallfiles.avgsize` 設定未満の場合にのみトリガされます。最終出力パスが HDFS 内にある場合は、Amazon EMR Hive でも動作はまったく同じです。出力パスが Amazon S3 内の場合、`hive.merge.smallfiles.avgsize` パラメータは無視されます。その場合、`hive.merge.mapfiles` が `true` に設定されているとマージタスクが常にトリガされます。

#### ACID トランザクションおよび Amazon S3

Amazon EMR 6.1.0 以降では、Hive ACID (不可分性、整合性、分離性、耐久性) トランザクションをサポートしているため、データベースの ACID プロパティに準拠しています。この機能を使用すると、Amazon Simple Storage Service (Amazon S3) のデータを使用して Hive マネージドテーブルで INSERT、UPDATE、DELETE および MERGE の各オペレーションを実行できます。

#### Hive Live Long and Process (LLAP)

デフォルトの Apache Hive のバージョン 2.0 で追加された [LLAP 機能](#) は、Amazon EMR リリース 5.0 の Hive 2.1.0 ではサポートされません。

Amazon EMR バージョン 6.0.0 以降では、Hive の Live Long and Process (LLAP) 機能をサポートしています。詳細については、「[Hive LLAP の使用](#)」を参照してください。

## Amazon EMR リリースバージョン 4.x と 5.x の Hive の相違点

このセクションでは、Amazon EMR リリース 4.x 上の Hive バージョン 1.0.0 から Amazon EMR リリース 5.x 上の Hive 2.x に Hive の実装を移行する前に考慮すべき相違点について説明します。

### 運用面での相違点と考慮事項

- [ACID \(不可分性、整合性、分離性、耐久性\) トランザクション](#)のサポートの追加: Amazon EMR 4.x 上の Hive 1.0.0 とデフォルトの Apache Hive のこの相違点は排除されています。
- Amazon S3 への直接書き込みの排除: Amazon EMR 上の Hive 1.0.0 とデフォルトの Apache Hive のこの相違点は排除されています。Amazon EMR リリース 5.x 上の Hive 2.1.0 は、現在、Amazon S3 に保存される一時ファイルの作成、読み取り、書き込みを行います。結果として、同じ表の読み取りと書き込みのために、対処法としてクラスターのローカル HDFS ファイルシステムに一時テーブルを作成する必要はなくなります。バージョン対応バケットを使用する場合、これらの一時ファイルを必ず以下のように管理してください。
- Amazon S3 バージョニング対応バケットを使用する場合の一時ファイルの管理: 生成されたデータの出力先が Amazon S3 である Hive クエリを実行すると、多くの一時ファイルとディレクトリが作成されます。これは、前述の新しい動作です。バージョン対応 S3 バケットを使用する場合、これらの一時ファイルのために Amazon S3 が煩雑になり、削除されないならコストが発生します。/\_tmp プレフィックスが付けられたデータが 5 日などの短期間で削除されるように、ライフサイクルルールを調整します。詳細については、「[ライフサイクル設定の指定](#)」を参照してください。
- Log4j は Log4j 2 に更新: Log4j を使用する場合、このアップグレードのためにログ設定を変更しなければならない場合があります。詳細については、[Apache log4j 2](#) を参照してください。

### パフォーマンスの相違点と考慮事項

- Tez とのパフォーマンスの違い: Amazon EMR リリース 5.x では、Tez はではなく Hive のデフォルトの実行エンジンです MapReduce。Tez は、ほとんどのワークフローでパフォーマンスを向上させます。
- 多くのパーティションを持つテーブル: 多数の動的パーティションを生成するクエリは失敗することがあり、多くのパーティションを持つテーブルから選択するクエリは予想されるより時間がかかる場合があります。たとえば、100,000 個のパーティションからの選択には 10 分以上かかる場合があります。

## Amazon EMR での Hive の追加機能

Amazon EMR は、Amazon Simple Storage Service (Amazon S3) や DynamoDB との間で読み書きする機能など、他の AWS サービスとの Hive 統合をサポートする新機能で Hive を拡張します。

### Hive の変数

Hive の引数スクリプト内で変数を使用するには、ドル記号と中括弧を使用します。

```
add jar ${LIB}/jsonserde.jar
```

これらの変数の値を Hive に渡すには、次の例で示すようにコマンドラインで `-d` パラメータを使用します。

```
-d LIB=s3://elasticmapreduce/samples/hive-ads/lib
```

また、Hive スクリプトを実行するステップに値を渡すこともできます。

コンソールを使用して Hive のステップに変数値を渡すには

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [クラスターを作成] を選択します。
3. [Steps] セクションの [Add Step] で、リストから [Hive Program] を選択し、[Configure and add] を選択します。
4. [Add Step] ダイアログで、次の表を参考にしながらパラメータを指定し、[Add] を選択します。

フィールド	アクション
スクリプト S3 場所*	Amazon S3 でスクリプトがある場所の URI を指定します。値は <i>BucketName /path /</i> の形式である必要があります <i>ScriptName</i> 。 例: <code>s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q</code> 。
S3 の場所の入力	Amazon S3 で入力ファイルがある場所の URI を指定します (オプション)。値は <i>BucketName /path /</i> の形式である必要があります。指定する


フィールド	アクション
	と、この値は INPUT という名前のパラメータとして Hive スクリプトに渡されます。例: s3://elasticmapreduce/samples/hive-ads/tables/ 。
S3 の場所の出力	Amazon S3 で出力を保存する場所の URI を指定します (オプション)。値は <i>BucketName /path</i> の形式である必要があります。指定すると、この値は OUTPUT という名前のパラメータとして Hive スクリプトに渡されます。例: s3://mybucket/hive-ads/output/ 。
引数	<p>引数のリスト (スペース区切りの文字列) を入力し、Hive に渡します (オプション)。\${SAMPLE} という名前の Hive スクリプトでパス変数を定義した場合の例: )</p> <pre data-bbox="451 779 1507 1016">CREATE EXTERNAL TABLE logs (requestBeginTime STRING, requestEndTime STRING, hostname STRING) PARTITIONED BY (dt STRING) \ ROW FORMAT serde 'com.amazon.elasticmapreduce.JsonSerde' WITH SERDEPROPERTIES ( 'paths'='requestBeginTime, requestEndTime, hostname' ) LOCATION '\${SAMPLE}/tables/impressions';</pre> <p>変数の値を渡すには、[Arguments] ウィンドウに次のように入力します。</p> <pre data-bbox="451 1136 1507 1171">-d SAMPLE=s3://elasticmapreduce/samples/hive-ads/ 。</pre>
Action on Failure	<p>エラーに反応しているクラスターを判断します。この設定に使用できる値は次のとおりです。</p> <ul data-bbox="451 1346 1507 1682" style="list-style-type: none"> <li>• Terminate cluster: ステップが失敗した場合、クラスターを停止します。クラスターの停止保護が有効で、キープアライブが有効な場合は、クラスターは停止されません。</li> <li>• Cancel and wait: ステップが失敗した場合、残りのステップをキャンセルします。クラスターのキープアライブが有効な場合、クラスターは停止されません。</li> <li>• Continue: ステップが失敗した場合、次のステップに進みます。</li> </ul>

5. 必要に応じて値を選択し、[Create cluster] を選択します。



を使用して Hive ステップに変数値を渡すには AWS CLI

を使用して Hive ステップに変数値を渡すには AWS CLI、`--steps`パラメータを使用して引数リストを含めます。

 Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.1.0 \  
--applications Name=Hive Name=Pig --use-default-roles --ec2-attributes \  
  KeyName=myKey --instance-type m5.xlarge --instance-count 3 \  
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,Args=[-f,s3:// \  
elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,INPUT=s3:// \  
elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://mybucket/hive-ads/output/, \  
-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

での Amazon EMR コマンドの使用の詳細については AWS CLI、「」を参照してください <https://docs.aws.amazon.com/cli/latest/reference/emr>。

Java SDK を使用して Hive のステップに変数値を渡すには

- 次の例は、SDK を使用してステップに変数を渡す方法を示しています。詳細については、AWS SDK for Java API リファレンスの「[Class StepFactory](#)」を参照してください。

```
StepFactory stepFactory = new StepFactory();  
  
StepConfig runHive = new StepConfig()  
  .withName("Run Hive Script")  
  .withActionOnFailure("TERMINATE_JOB_FLOW")  
  .withHadoopJarStep(stepFactory.newRunHiveScriptStep("s3://mybucket/script.q",  
    Lists.newArrayList("-d","LIB= s3://elasticmapreduce/samples/hive-ads/lib"));
```

## 部分 DynamoDB スキーマに対応するための Amazon EMR Hive クエリ

Amazon EMR Hive は、DynamoDB テーブルをクエリすると、すべての列をクエリに含める必要はなく、データをフィルターできる列のサブセットを指定できるので、柔軟性が最大になります。この部分スキーマクエリの手法は、スパースデータベーススキーマを使用し、タイムスタンプでのフィルタのように少数の列に基づいてレコードをフィルタする場合に効果的です。

以下の例では、Hive クエリの使用方法を示します。

- DynamoDB テーブルを作成する。
- DynamoDB で項目 (行) のサブセットを選択し、データを特定の列までさらに絞り込む。
- 結果のデータを Amazon S3 にコピーする。

```
DROP TABLE dynamodb;
DROP TABLE s3;

CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, fullColumn
map<String, String>)
  STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
  TBLPROPERTIES (
    "dynamodb.table.name" = "myTable",
    "dynamodb.throughput.read.percent" = ".1000",
    "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");

CREATE EXTERNAL TABLE s3(map<String, String>)
  ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
  LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE s3 SELECT item fullColumn FROM dynamodb WHERE recordTimeStamp <
"2012-01-01";
```

次の表では、DynamoDB から項目の任意の組み合わせを選択するためのクエリ構文を示します。

クエリの例	結果の説明
<code>SELECT * FROM <i>table_name</i> ;</code>	特定のテーブルからすべての項目 (行) を選択し、それらの項目で使用できるすべての列のデータを含めます。

クエリの例	結果の説明
<code>SELECT * FROM <i>table_name</i> WHERE <i>field_name</i> =<i>value</i>;</code>	特定のテーブルから一部の項目（行）を選択し、それらの項目で利用できるすべての列のデータを含めます。
<code>SELECT <i>column1_name</i> , <i>column2_name</i> , <i>column3_name</i> FROM <i>table_name</i> ;</code>	特定のテーブルからすべての項目（行）を選択し、それらの項目で利用できる一部の列のデータを含めます。
<code>SELECT <i>column1_name</i> , <i>column2_name</i> , <i>column3_name</i> FROM <i>table_name</i> WHERE <i>field_name</i> =<i>value</i>;</code>	特定のテーブルから一部の項目（行）を選択し、それらの項目で利用できる一部の列のデータを含めます。

## 異なる AWS リージョン内の DynamoDB テーブル間でデータをコピーする

Amazon EMR Hive には、DynamoDB テーブルごとに設定できる `dynamodb.region` プロパティがあります。2 つのテーブルで `dynamodb.region` の設定が異なると、指定されたリージョンの間でテーブル間のデータのコピーが自動的に行われます。

次の例では、`dynamodb.region` プロパティが設定された Hive スクリプトを含む DynamoDB テーブルの作成方法を示します。

### Note

テーブルリージョンごとのプロパティは、グローバル Hive プロパティをオーバーライドします。

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String, String> fullColumn)
  STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
  TBLPROPERTIES (
    "dynamodb.table.name" = "myTable",
    "dynamodb.region" = "eu-west-1",
    "dynamodb.throughput.read.percent" = ".1000",
    "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

## テーブルごとに DynamoDB スループット値を設定する

Amazon EMR Hive では、テーブル定義でテーブルごとに DynamoDB readThroughputPercent と writeThroughputPercent の設定を設定できます。以下の Amazon EMR Hive スクリプトでは、スループット値の設定方法を示します。DynamoDB スループット値の詳細については、「[Specifying Read and Write Requirements for Tables](#)」を参照してください。

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String, String> fullColumn)
  STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
  TBLPROPERTIES (
    "dynamodb.table.name" = "myTable",
    "dynamodb.throughput.read.percent" = ".4",
    "dynamodb.throughput.write.percent" = "1.0",
    "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

## Hive の外部メタストアの設定

Hive のメタストア情報の記録先は、デフォルトでは、プライマリノードのファイルシステムにある MySQL データベースです。メタストアには、テーブルとテーブル構築の基となるデータ (パーティション名やデータタイプなど) が含まれています。クラスターが終了すると、プライマリノードも含めて、すべてのクラスターノードがシャットダウンされます。この場合、ノードファイルシステムではエフェメラルストレージが使用されているため、ローカルデータが失われます。メタストアを永続的に保存する場合は、クラスター外に外部メタストアを作成する必要があります。

外部メタストアとして 2 つのオプションがあります。

- AWS Glue Data Catalog (Amazon EMR リリース 5.8.0 以降のみ)。

詳細については、「[AWS Glue データカタログを Hive のメタストアとして使用する](#)」を参照してください。

- Amazon RDS または Amazon Aurora。

詳細については、「[外部の MySQL データベースまたは Amazon Aurora の使用](#)」を参照してください。

**Note**

Hive 3 を使用していて、Hive メタストアへの接続が多すぎる場合は、パラメータ `datanucleus.connectionPool.maxPoolSize` の値を小さくするか、データベースサーバーが処理できる接続数を増やします。接続数が増えるのは、Hive が JDBC 接続の最大数を計算する方法によるものです。パフォーマンスの最適な値を計算するには、「[Hive Configuration Properties](#)」を参照してください。

## AWS Glue データカタログを Hive のメタストアとして使用する

Amazon EMR リリース 5.8.0 以降を使用して、AWS Glue Data Catalog をメタストアとして使用するように Hive を設定できます。永続的なメタストア、またはさまざまなクラスター、サービス、アプリケーション、あるいは AWS アカウントで共有されるメタストアが必要である場合は、この設定をお勧めします。

AWS Glue はフルマネージド型の抽出、変換、ロード (ETL) サービスで、データの分類、クリーニング、強化、さまざまなデータストア間の信頼性の高い移動を簡単かつ費用対効果の高い方法で実現します。AWS Glue Data Catalog は、Amazon EMR だけでなく、Amazon RDS、Amazon Redshift、Redshift Spectrum、Athena、および Apache Hive メタストアと互換性のあるすべてのアプリケーションと統合されたメタデータリポジトリを提供します。AWS Glue クローラは、Amazon S3 のソースデータからスキーマを自動的に推測し、関連するメタデータをデータカタログに保存できます。データカタログの詳細については、「[AWS Glue デベロッパーガイド](#)」の「[Glue データカタログの入力](#)」を参照してください。AWS

AWS Glue には個別の料金が適用されます。Data Catalog にメタデータを保存してアクセスするための月額料金、AWS Glue ETL ジョブとクローラランタイムに対して 1 分あたりに課金される時間料金、プロビジョニングされた開発エンドポイントごとに 1 分あたりに課金される時間料金があります。Data Catalog では、最大 100 万個までのオブジェクトを無料で保存できます。100 万を超えるオブジェクトを保存した場合は、100,000 オブジェクトごとに 1 USD が課金されます。Data Catalog 内のオブジェクトは、テーブル、パーティション、またはデータベースです。詳細については、「[Glue 料金表](#)」を参照してください。

**Important**

2017 年 8 月 14 日より前に Amazon Athena または Amazon Redshift Spectrum を使用してテーブルを作成した場合、データベースとテーブルは AWS Glue データカタログとは別の Athena 管理カタログに保存されます。Amazon EMR をこれらのテーブルと統合する

には、AWS Glue データカタログにアップグレードする必要があります。詳細については、[「Amazon Athena ユーザーガイド」の「AWS Glue データカタログへのアップグレード」](#)を参照してください。Amazon Athena

## AWS Glue Data Catalog をメタストアとして指定する

、または Amazon EMR API を使用して AWS Management Console AWS CLI、AWS Glue データカタログをメタストアとして指定できます。CLI または API を使用する場合は、Hive の設定分類を使用して Data Catalog を指定します。さらに、Amazon EMR 5.16.0 以降では、設定分類を使用して、別のデータカタログを指定できます AWS アカウント。コンソールを使用する場合は、[Advanced Options] (詳細オプション) または [Quick Options] (クイックオプション) を使用して、Data Catalog を指定できます。

### New console

新しいコンソールで AWS Glue Data Catalog を Hive メタストアとして指定するには

1. にサインインし AWS Management Console、<https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. 左側のナビゲーションペインの [EMR on EC2] で、[クラスター] を選択し、[クラスターの作成] を選択します
3. [アプリケーションバンドル] で [Core Hadoop]、[HBase]、または [カスタム] を選択します。クラスターをカスタマイズする場合は、必ず Hive または HCatalog をアプリケーションの 1 つとして選択しします。
4. [AWS Glue Data Catalog 設定] で [Hive テーブルメタデータに使用] チェックボックスをオンにします。
5. クラスターに適用するその他のオプションを選択します。
6. クラスターを起動するには、[クラスターの作成] を選択します。

### Old console

古いコンソールで AWS Glue Data Catalog を Hive メタストアとして指定するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。

2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Release] で、[emr-5.8.0] 以降を選択します。
4. [Release (リリース)] で [Hive] または [HCatalog] を選択します。
5. [AWS Glue Data Catalog 設定] で [Hive テーブルメタデータに使用] を選択します。
6. 必要に応じてクラスターの他のオプションを選択し、[Next (次へ)] 選択してアプリケーションに適切な他のクラスターオプションを設定します。

## CLI

を使用して AWS Glue データカタログを Hive メタストアとして指定するには AWS CLI

AWS CLI と EMR API を使用して設定分類を指定する方法の詳細については、「」を参照してください [アプリケーションの設定](#)。

- 次の例に示すように、hive-site 設定分類を使用して `hive.metastore.client.factory.class` の値を指定します。

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
      "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

EMR リリースバージョン 5.28.0、5.28.1、5.29.0、または 6.x で、AWS Glue Data Catalog をメタストアとして使用してクラスターを作成する場合は、`hive.metastore.schema.validation` を `false` に設定します。これにより、Hive と HCatalog が MySQL に対してメタストアスキーマを検証できなくなります。この設定がない場合、プライマリインスタンスグループは Hive または HCatalog の再設定後に中断されません。

```
[
  {
```

```
"Classification": "hive-site",
"Properties": {
  "hive.metastore.client.factory.class":
"com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
  "hive.metastore.schema.verification": "false"
}
}
]
```

EMR リリースバージョン 5.28.0、5.28.1、または 5.29.0 にクラスターが既に存在する場合は、次の情報を使用してプライマリインスタンスグループ `hive.metastore.schema.verification` を `false` に設定できます。

```
Classification = hive-site
Property       = hive.metastore.schema.verification
Value         = false
```

別の AWS アカウントでデータカタログを指定するには、次の例に示すように `hive.metastore.glue.catalogid` プロパティを追加します。`acct-id` は、データカタログの AWS アカウントに置き換えます。

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
"com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```



## IAM アクセス許可

クラスターの EC2 インスタンスプロファイルには、AWS Glue アクションの IAM アクセス許可が必要です。さらに、AWS Glue Data Catalog オブジェクトの暗号化を有効にする場合は、暗号化 AWS KMS key に使用される の暗号化、復号、生成もロールに許可する必要があります。

### AWS Glue アクションのアクセス許可

Amazon EMR のデフォルトの EC2 インスタンスプロファイルを使用する場合、アクションは必要ありません。にアタッチされている AmazonElasticMapReduceforEC2Role マネージドポリシーは、必要なすべての AWS Glue アクション EMR\_EC2\_DefaultRole を許可します。ただし、カスタム EC2 インスタンスプロファイルとアクセス許可を指定する場合は、適切な AWS Glue アクションを設定する必要があります。開始点として AmazonElasticMapReduceforEC2Role 管理ポリシーを使用します。詳細については、「Amazon EMR 管理ガイド」の「[クラスター EC2 インスタンスのサービスロール \(EC2 インスタンスプロファイル\)](#)」を参照してください。

### AWS Glue Data Catalog を暗号化および復号するためのアクセス許可

インスタンスプロファイルには、キーを使用してデータを暗号化および復号するためのアクセス許可が必要です。以下のステートメントが両方とも適用される場合、これらのアクセス許可を設定する必要はありません。

- AWS Glue のマネージドキーを使用して、AWS Glue Data Catalog オブジェクトの暗号化を有効にします。
- AWS Glue データカタログ AWS アカウント と同じ にあるクラスターを使用します。

それ以外の場合は、EC2 インスタンスプロファイルにアタッチされたアクセス許可ポリシーに次のステートメントを追加する必要があります。

```
[
  {
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Action": [
          "kms:Decrypt",
          "kms:Encrypt",
          "kms:GenerateDataKey"
        ]
      }
    ]
  }
]
```

```
        "Resource": "arn:aws:kms:region:acct-id:key/12345678-1234-1234-1234-123456789012"
      }
    ]
  }
]
```

AWS Glue データカタログの暗号化の詳細については、「[AWS Glue デベロッパーガイド](#)」の「[データカタログの暗号化](#)」を参照してください。

### リソースベースのアクセス許可

Amazon EMR で AWS Glue を Hive、Spark、または Presto と組み合わせて使用する場合、AWS Glue は Data Catalog リソースへのアクセスを制御するためのリソースベースのポリシーをサポートします。これらのリソースには、データベース、テーブル、接続、ユーザー定義関数が含まれます。詳細については、「[AWS Glue デベロッパーガイド](#)」の「[AWS Glue リソースポリシー](#)」を参照してください。

リソースベースのポリシーを使用して Amazon EMR 内から AWS Glue へのアクセスを制限する場合、アクセス許可ポリシーで指定するプリンシパルは、クラスターの作成時に指定する EC2 インスタンスプロファイルに関連付けられたロール ARN である必要があります。例えば、カタログにタッチされたリソースベースのポリシーの場合、次の例に示す形式を使用して、クラスター EC2 インスタンスのデフォルトのサービスロールのロール ARN `Principal`、`EMR_EC2_DefaultRole` をとして指定できます。

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

`acct-id` は AWS Glue アカウント ID とは異なる場合があります。これにより、さまざまなアカウントで EMR クラスターからアクセスできます。異なるアカウントから、複数のプリンシパルを指定できます。

### AWS Glue Data Catalog を使用する場合の考慮事項

AWS Glue データカタログを Hive のメタストアとして使用する場合は、次の項目を考慮してください。

- Hive シェルを使用した補助 JAR の追加はサポートされません。この問題を回避するには、hive-site 設定分類を使用して `hive.aux.jars.path` プロパティを設定します。これにより、Hive のクラスパスに補助 JAR が追加されます。
- [Hive のトランザクション](#) はサポートされていません。

- AWS Glue 内からのテーブルの名前変更はサポートされていません。
- LOCATION を指定せずに Hive テーブルを作成すると、テーブルデータは、hive.metastore.warehouse.dir プロパティによって指定された場所に保管されます。デフォルトでは、これは HDFS 内の場所です。別のクラスターがテーブルにアクセスする必要がある場合、テーブルを作成したクラスターに対する適切なアクセス許可がない限り、処理に失敗します。さらに、HDFS ストレージは一時的であるため、クラスターが終了すると、テーブルデータは失われ、テーブルを再作成する必要があります。AWS Glue を使用して Hive Amazon S3 LOCATION を指定することをお勧めします。または、hive-site 設定分類を使用して、hive.metastore.warehouse.dir の Amazon S3 で場所を指定できます。これは、すべての Hive テーブルに適用されます。テーブルが HDFS ロケーションで作成され、テーブルを作成したクラスターがまだ実行されている場合は、AWS Glue 内からテーブルロケーションを Amazon S3 に更新できます。詳細については、「[AWS Glue デベロッパーガイド](#)」の「[Glue コンソールでのテーブルの操作](#)」を参照してください。AWS
- 引用符とアポストロフィを含むパーティション値はサポートされていません(例: PARTITION (owner="Doe's")).
- [列統計](#)は、emr-5.31.0 以降でサポートされています。
- [Hive 認可](#)の使用はサポートされていません。代替策として、[AWS Glue リソースベースのポリシー](#)を使用することを検討してください。詳細については、[AWS 「Glue データカタログへの Amazon EMR アクセスにリソースベースのポリシーを使用する」](#)を参照してください。
- [Hive 制約](#)はサポートされていません。
- [Hive でのコストベースの最適化](#)はサポートされていません。
- hive.metastore.partition.inherit.table.properties の使用はサポートされていません。
- 次のメタストア定数の使用はサポートされていません。BUCKET\_COUNT, BUCKET\_FIELD\_NAME, DDL\_TIME, FIELD\_TO\_DIMENSION, FILE\_INPUT\_FORMAT, FILE\_OUTPUT\_FORMAT, HIVE\_FILTER\_FIELD\_LAST\_ACCESS, HIVE\_FILTER\_FIELD\_OWNER, HIVE\_FILTER\_FIELD\_PARAMS, IS\_ARCHIVED, META\_TABLE\_COLUMNS, META\_TABLE\_COLUMN\_TYPES, META\_TABLE\_DB, META\_TABLE\_LOCATION, META\_TABLE\_NAME, META\_TABLE\_PARTITION\_COLUMNS, META\_TABLE\_SERDE, META\_TABLE\_STORAGE, ORIGINAL\_LOCATION。
- 述語式を使用する場合、明示的な値が比較演算子の右側にある必要があります。そうしなければ、クエリが失敗する可能性があります。
  - 正: SELECT \* FROM mytable WHERE time > 11
  - 誤: SELECT \* FROM mytable WHERE 11 > time

- Amazon EMR バージョン 5.32.0 および 6.3.0 以降では、述語式でのユーザー定義関数 (UDF) の使用がサポートされています。以前のバージョンを使用している場合、Hive がクエリの実行を最適化しようとするため、クエリが失敗する可能性があります。
- [一時テーブル](#) はサポートされていません。
- AWS Glue を使用して直接作成するのではなく、Amazon EMR を使用してアプリケーションを使用してテーブルを作成することをお勧めします。AWS Glue を使用してテーブルを作成すると、必須フィールドが欠落し、クエリ例外が発生する可能性があります。
- EMR 5.20.0 以降では、AWS Glue Data Catalog をメタストアとして使用すると、Spark と Hive の並列パーティションプルーニングが自動的に有効になります。この変更により、複数のリクエストを並列に実行してパーティションを取得できるため、クエリ計画時間が大幅に短縮されます。同時に実行できるセグメントの総数は、1~10 の範囲です。デフォルト値は 5 であり、これが推奨される設定です。hive-site 設定分類の `aws.glue.partition.num.segments` プロパティを指定すると、これを変更できます。スロットリングが発生した場合は、値を 1 に変更して機能をオフにすることができます。詳細については、「[AWS Glue セグメント構造](#)」を参照してください。

## 外部の MySQL データベースまたは Amazon Aurora の使用

Hive メタストアとして外部の MySQL データベースまたは Amazon Aurora を使用するには、Hive のメタストアのデフォルト設定値を上書きし、外部のデータベースの場所 (Amazon RDS MySQL インスタンスまたは Amazon Aurora PostgreSQL インスタンス) を指定します。

### Note

Hive では、メタストアテーブルに対する同時書き込みアクセスのサポートも防止も行いません。2 つのクラスター間でメタストア情報を共有する場合は、同じメタストアテーブルに同時に書き込まないようにします (同じメタストアテーブルの異なるパーティションに書き込む場合は除きます)。

次の手順では、Hive メタストアの場所として設定したデフォルト値を上書きし、新しく設定したメタストアの場所を使用してクラスターを起動する方法を示します。

EMR クラスターの外にメタストアを作成するには

1. MySQL または Aurora PostgreSQL データベースを作成します。PostgreSQL を使用する場合は、クラスターをプロビジョニングした後に設定する必要があります。クラスターの作成では

MySQL のみがサポートされます。Aurora MySQL と Aurora PostgreSQL の違いについては、「[Amazon Aurora MySQL の概要](#)」および「[Amazon Aurora PostgreSQL の操作](#)」を参照してください。Amazon RDS データベースの作成方法については、<https://aws.amazon.com/rds/> を参照してください。

- データベースと ElasticMapReduce-Master セキュリティグループ間の JDBC 接続を許可するようにセキュリティグループを変更します。アクセス用のセキュリティグループの変更方法については、「[Amazon EMR マネージドセキュリティグループの使用](#)」を参照してください。
- JDBC 設定値を、hive-site.xml に設定します:

#### Important

パスワードなどの機密情報を Amazon EMR 設定 API に渡す場合、その情報は、十分なアクセス許可を持つアカウントに表示されます。その情報が他のユーザーに表示される可能性のあることを懸念する場合は、管理者アカウントでクラスターを作成し、elasticmapreduce:DescribeCluster API キーに対するアクセス権限を明示的に拒否するロールを作成することで、他のユーザー (IAM ユーザーまたは認証情報を委譲されたユーザー) に対してそのクラスター上のサービスへのアクセスを制限します。

- 次の例のように、hiveConfiguration.json の編集を含む hive-site.xml という設定ファイルを作成します。

*hostname* を、データベースを実行する Amazon RDS インスタンスの DNS アドレスに置き換え、*username* と *password* を、データベースの認証情報に置き換えます。MySQL データベースインスタンスおよび Aurora データベースインスタンスへの接続の詳細については、「Amazon RDS ユーザーガイド」の「[MySQL データベースエンジンを実行している DB インスタンスへの接続](#)」および「[Athena DB クラスターへの接続](#)」を参照してください。javax.jdo.option.ConnectionURL は、JDBC メタストアの JDBC 接続文字列です。javax.jdo.option.ConnectionDriverName は、JDBC メタストアのドライバークラス名です。

MySQL JDBC ドライバーは Amazon EMR によってインストールされます。

値のプロパティには、スペースや改行を含めることはできません。すべてを 1 行にする必要があります。

```
[  
  {
```

```

    "Classification": "hive-site",
    "Properties": {
      "javax.jdo.option.ConnectionURL": "jdbc:mysql://hostname:3306/hive?
createDatabaseIfNotExist=true",
      "javax.jdo.option.ConnectionDriverName": "org.mariadb.jdbc.Driver",
      "javax.jdo.option.ConnectionUserName": "username",
      "javax.jdo.option.ConnectionPassword": "password"
    }
  }
]

```

- b. 次の AWS CLI コマンドに示すように、クラスターの作成時に `hiveConfiguration.json` ファイルを参照します。このコマンドでは、ファイルはローカルに保存されます。ファイルを Amazon S3 にアップロードし、そこで参照することもできます (例: `s3://DOC-EXAMPLE-BUCKET/hiveConfiguration.json`)。

**Note**

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```

aws emr create-cluster --release-label emr-7.1.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=Hive --configurations file://hiveConfiguration.json --use-
default-roles

```

4. クラスターのプライマリノードに接続します。

プライマリノードに接続する方法の詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してプライマリノードに接続する](#)」を参照してください。

5. 次のようなコマンドを入力することにより、Amazon S3 上の場所を指定して Hive テーブルを作成します。

```

CREATE EXTERNAL TABLE IF NOT EXISTS table_name
(
  key int,
  value int
)

```



```
LOCATION s3://DOC-EXAMPLE-BUCKET/hdfs/
```

6. 実行中のクラスターに Hive スクリプトを追加します。

Hive クラスターは、Amazon RDS に配置されたメタストアを使って実行されます。このメタストアを共有する他の Hive クラスターはすべて、メタストアの場所を指定して起動します。

## Hive JDBC ドライバーの使用

Microsoft Excel、MicroStrategy、QlikView、Tableau などの人気のビジネスインテリジェンスツールを Amazon EMR で使用して、データを調査しビジュアル化することができます。この種のツールの多くは Java Database Connectivity (JDBC) ドライバーまたは Open Database Connectivity (ODBC) ドライバーを必要とします。Amazon EMR は JDBC 接続と ODBC 接続の両方をサポートします。

以下の例では、SQL Workbench/J を SQL クライアントとして使用して、Amazon EMR の Hive クラスターに接続します。その他のドライバーについては、「[Amazon EMR でのビジネスインテリジェンスツールの使用](#)」を参照してください。

SQL Workbench/J をインストールして使用する前に、ドライバーパッケージをダウンロードしてドライバーをインストールします。パッケージに含まれているドライバーは、Amazon EMR リリースバージョン 4.0 以降で利用できる Hive バージョンをサポートしています。詳細なリリースノートおよびドキュメントについては、パッケージに含まれている PDF ドキュメントを参照してください。

- 最新の Hive JDBC ドライバーパッケージのダウンロード

<http://awssupportdatasvcs.com/bootstrap-actions/Simba/latest/>

- 古いバージョンの Hive JDBC ドライバー

<http://awssupportdatasvcs.com/bootstrap-actions/Simba/>

SQL Workbench をインストールして設定するには

1. オペレーティングシステム用の SQL Workbench/J クライアントを <http://www.sql-workbench.net/downloads.html> からダウンロードします。
2. SQL Workbench/J をインストールします。詳細については、SQL Workbench/J ユーザーズマニュアルの「[Installing and starting SQL Workbench/J](#)」を参照してください。
3. Linux, Unix, Mac OS X ユーザー: ターミナルセッションで、以下のコマンドを使用して、クラスターのマスターノードへの SSH トンネルを作成します。*master-public-dns-name* をマ

ターノードのパブリック DNS 名に置き換え、*path-to-key-file* を Amazon EC2 プライベートキー (.pem) ファイルの場所とファイル名に置き換えます。

```
ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000
hadoop@master-public-dns-name
```

Windows ユーザー: PuTTY セッションで、ソースポートを 10000、送信先を *master-public-dns-name*:10000 としてクラスターのマスターノードへの SSH トンネルを作成します (ローカルポートフォワーディングを使用)。*master-public-dns-name* をマスターノードのパブリック DNS 名に置き換えます。

4. SQL Workbench に JDBC ドライバを追加します。
  - a. [Select Connection Profile (接続プロファイルの選択)] ダイアログボックスで [Manage Drivers (ドライバーの管理)] をクリックします。
  - b. [Create a new entry (新しいエントリの作成)] (空白ページ) アイコンをクリックします。
  - c. 名前フィールドに、**Hive JDBC** と入力します。
  - d. [Library] で [Select the JAR file(s)] アイコンをクリックします。
  - e. 抽出したドライバーがある場所へ移動します。ダウンロードした JDBC ドライバーパッケージバージョンに含まれているドライバーを選択し、[開く] をクリックします。

例えば、JDBC ドライバーパッケージには以下の JAR が含まれている場合があります。

```
hive_metastore.jar
hive_service.jar
HiveJDBC41.jar
libfb303-0.9.0.jar
libthrift-0.9.0.jar
log4j-1.2.14.jar
ql.jar
slf4j-api-1.5.11.jar
slf4j-log4j12-1.5.11.jar
TCLIServiceClient.jar
zookeeper-3.4.6.jar
```

- f. [Please select one driver] (ドライバーを 1 つ選択してください) ダイアログボックスで、`com.amazon.hive.jdbc41.HS2Driver` を選択して、[OK] を選択します。
5. [Manage Drivers] (ドライバーの管理) ダイアログボックスに戻ったら、[Classname] (クラス名) フィールドに値が設定されていることを確認し、[OK] を選択します。



6. [Select Connection Profile] (接続プロファイルの選択) ダイアログボックスに戻ったら、[Driver] (ドライバー) フィールドが [Hive JDBC] に設定されていることを確認し、[URL] フィールドに JDBC 接続文字列 `jdbc:hive2://localhost:10000/default` を入力します。
7. [OK] を選択すると接続します。接続が完了すると、接続の詳細が SQL Workbench/J のウィンドウの上部に表示されます。

Hive と JDBC インターフェイスの使用方法の詳細については、Apache Hive のドキュメントの「[HiveClient](#)」および「[HiveJDBCInterface](#)」を参照してください。

## Hive のパフォーマンスの向上

Amazon EMR には、Hive を使用して Amazon S3 に保存されたデータのクエリ、読み取り、および書き込みを行うときのパフォーマンスを最適化するのに役立つ機能が用意されています。

S3 Select では、Amazon S3 に処理を「プッシュダウン」することで一部のアプリケーションの CSV および JSON ファイルのクエリパフォーマンスを向上させることができます。

EMRFS S3 最適化コミッターは、[OutputCommitter](#) クラスの代替手段であり、EMRFS を使用して Amazon S3 ファイルを書き込む際のパフォーマンスを向上させるために、リスト操作と名前変更操作を排除します。

### トピック

- [Hive EMRFS S3 向けに最適化されたコミッターを有効にする](#)
- [S3 Select と Hive を使用してパフォーマンスを向上させる](#)
- [MSCK の最適化](#)

## Hive EMRFS S3 向けに最適化されたコミッターを有効にする

Hive EMRFS S3 向けに最適化されたコミッターは、EMRFS を使用するとき EMR Hive が挿入クエリ用のファイルを書き込む代替方法です。コミッターは Amazon S3 で行うリスト操作と名前変更操作を不要にし、アプリケーションのパフォーマンスを向上させます。この機能は EMR 5.34 および EMR 6.5 以降で利用可能です。

## コミッターを有効にする

EMR Hive が HiveEMRFSOptimizedCommitter を使用してすべての Hive 管理テーブルと外部テーブルのデフォルトとしてデータをコミットできるようにするには、EMR 6.5.0 または EMR 5.34.0 クラスターで次の hive-site 設定を使用してください。

```
[
  {
    "classification": "hive-site",
    "properties": {
      "hive.blobstore.use.output-committer": "true"
    }
  }
]
```

### Note

hive.exec.parallel が true に設定されているときは、この機能をオンにしないでください。

## 制限事項

タグには以下のベーシックな制限があります。

- Hive で小さなファイルを自動的にマージできるようにすることはサポートされていません。最適化されたコミッターが有効になっている場合でも、デフォルトの Hive コミットロジックが使用されます。
- Hive ACID テーブルはサポートされていません。最適化されたコミッターが有効になっている場合でも、デフォルトの Hive コミットロジックが使用されます。
- 書き込まれるファイルのファイル命名法が Hive の <task\_id>\_<attempt\_id>\_<copy\_n> から <task\_id>\_<attempt\_id>\_<copy\_n>\_<query\_id> に変更されました。例えば、次の名前のファイルがあるとします。

```
s3://warehouse/table/partition=1/000000_0 は s3://warehouse/table/
partition=1/000000_0-
hadoop_20210714130459_ba7c23ec-5695-4947-9d98-8a40ef759222-1 に変更されま
す。query_id はここでは、ユーザー名、タイムスタンプ、UUID を組み合わせたものです。
```

- カスタムパーティションが異なるファイルシステム (HDFS、S3) にある場合、この機能は自動的に無効になります。有効にすると、デフォルトの Hive コミットロジックが使用されます。

## S3 Select と Hive を使用してパフォーマンスを向上させる

Amazon EMR リリースバージョン 5.18.0 以降では、Amazon EMR での Hive を使用した [S3 Select](#) を使用できます。S3 Select では、アプリケーションはオブジェクトに含まれるデータのサブセットのみを取得できます。Amazon EMR では、大量のデータセットをフィルタリングして処理する計算作業をクラスターから Amazon S3 に「プッシュダウン」することにより、一部のアプリケーションのパフォーマンスを高めることができます。また、Amazon EMR と Amazon S3 の間で転送されるデータの量も削減されます。

S3 Select では、Hive セッション中に `s3select.filter` 設定変数を `true` に設定することで、CSV ファイルと JSON ファイルに基づいて Hive テーブルがサポートされています。詳細な説明と例については、「[コードで S3 Select を指定する](#)」を参照してください。

### S3 Select が使用するアプリケーションに適しているかどうかを確認する

S3 Select が使用するアプリケーションに適しているかどうかを確認するために、S3 Select を使用した場合と使用しない場合のアプリケーションのベンチマークを行うことをお勧めします。

アプリケーションが S3 Select を使用する候補となるかどうかを判断するには、次のガイドラインを使用します。

- クエリは元のデータセットの半分以上を除外する。
- クエリのフィルター述語では、Amazon S3 Select がサポートするデータ型を持つ列が使用される。詳細については、「Amazon Simple Storage Service ユーザーガイド」の「[データ型](#)」を参照してください。
- Amazon S3 と Amazon EMR クラスター間のネットワーク接続は、転送速度と使用可能な帯域幅が良好です。Amazon S3 は、HTTP 応答を圧縮しないため、応答サイズは圧縮された入力ファイルと比較して増加する可能性があります。

### 考慮事項と制約事項

- お客様が用意した暗号化キーを使用した Amazon S3 サーバー側の暗号化 (SSE-C) とクライアント側の暗号化はサポートされていません。
- `AllowQuotedRecordDelimiters` プロパティはサポートされていません。このプロパティを指定した場合、クエリは失敗します。

- UTF-8 形式の CSV ファイルと JSON ファイルのみがサポートされています。複数行の CSV と JSON はサポートされません。
- 圧縮されていないファイルか gzip ファイルまたは bzip2 ファイルのみがサポートされます。
- 最後の行のコメント文字はサポートされていません。
- ファイルの末尾にある空の行は処理されません。
- Amazon EMR の Hive では、S3 Select がサポートするプリミティブデータ型がサポートされています。詳細については、「Amazon Simple Storage Service ユーザーガイド」の「[データ型](#)」を参照してください。

## コードで S3 Select を指定する

Hive テーブルで S3 Select を使用するには、INPUTFORMAT クラス名として `com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat` を指定してテーブルを作成し、TBLPROPERTIES 句を使用して `s3select.format` プロパティの値を指定します。

デフォルトでは、S3 Select はクエリの実行時に無効になっています。次に示すように Hive セッションで `s3select.filter` を `true` に設定して S3 Select を有効にします。以下の例では、基盤となる CSV ファイルと JSON ファイルからテーブルを作成するときに S3 Select を指定して、シンプルな SELECT ステートメントを使用してテーブルをクエリする方法を説明しています。

### Example CSV ベースのテーブルの CREATE TABLE ステートメント

```
CREATE TABLE mys3selecttable (  
  col1 string,  
  col2 int,  
  col3 boolean  
)  
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','  
STORED AS  
INPUTFORMAT  
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'  
OUTPUTFORMAT  
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'  
LOCATION 's3://path/to/mycsvfile/'  
TBLPROPERTIES (  
  "s3select.format" = "csv",  
  "s3select.headerInfo" = "ignore"  
);
```

## Example JSON ベースのテーブルの CREATE TABLE ステートメント

```
CREATE TABLE mys3selecttable (  
  col1 string,  
  col2 int,  
  col3 boolean  
)  
ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'  
STORED AS  
INPUTFORMAT  
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'  
OUTPUTFORMAT  
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'  
LOCATION 's3://path/to/json/'  
TBLPROPERTIES (  
  "s3select.format" = "json"  
);
```

## Example SELECT TABLE ステートメント

```
SET s3select.filter=true;  
SELECT * FROM mys3selecttable WHERE col2 > 10;
```

## MSCK の最適化

Hive は各テーブルのパーティションのリストをメタストアに保存します。ただし、パーティションがファイルシステムに直接追加されたり、ファイルシステムに削除されたりしても、Hive メタストアはこれらの変更を認識しません。[MSCK コマンド](#)は、ファイルシステムに直接追加または削除されたパーティションの Hive メタストア内のパーティションメタデータを更新します。コマンドの構文は次のとおりです。

```
MSCK [REPAIR] TABLE table_name [ADD/DROP/SYNC PARTITIONS];
```

Hive はこのコマンドを以下のように実装しています。

1. Hive はテーブルのすべてのパーティションをメタストアから取得します。次に、ファイルシステムに存在しないパーティションパスのリストから、メタストアから削除するパーティションのリストを作成します。

2. Hive はファイルシステムに存在するパーティションパスを収集し、メタストアにあるパーティションのリストと比較し、メタストアに追加する必要があるパーティションのリストを生成します。
3. Hive は、ADD、DROP、または SYNC モードを使用してメタストアを更新します。

#### Note

メタストアに多数のパーティションがある場合、ファイルシステムの exists API 呼び出しはパーティションごとに行う必要があるため、ファイルシステムにパーティションが存在しないかどうかを確認するステップの実行には時間がかかります。

Amazon EMR 6.5.0 では、Hive に `hive.emr.optimize.msck.fs.check` というフラグが導入されました。このフラグを有効にすると、Hive はファイルシステム API 呼び出しを行う代わりに、上記のステップ 2 で生成されたファイルシステムのパーティションパスのリストからパーティションの有無を確認します。Amazon EMR 6.8.0 では、Hive はこの最適化をデフォルトで有効にしたため、フラグ `hive.emr.optimize.msck.fs.check` を設定する必要がなくなりました。

## Hive Live Long and Process (LLAP) の使用

Amazon EMR 6.0.0 は、Hive の Live Long and Process (LLAP) 機能をサポートします。LLAP では、永続的なデーモンによるインテリジェントなメモリ内キャッシングを使用し、以前のデフォルトの Tez コンテナ実行モードと比較してクエリのパフォーマンスを向上させます。

Hive LLAP デーモンは、YARN サービスとして管理および実行されます。YARN サービスは長時間実行している YARN アプリケーションと見なすことができるため、一部のクラスターリソースは Hive LLAP 専用となり、他のワークロードには使用できません。詳細については、[LLAP](#) および [YARN サービス API](#) を参照してください。

## Amazon EMR で Hive LLAP を有効にする

Amazon EMR で Hive LLAP を有効にするには、クラスターを起動するときに次の設定を指定します。

```
[
  {
    "Classification": "hive",
    "Properties": {
```

```

    "hive.llap.enabled": "true"
  }
}
]

```

詳細については、「[アプリケーションの設定](#)」を参照してください。

デフォルトでは、Amazon EMR はクラスター YARN リソースの約 60% を Hive LLAP デーモンに割り当てます。Hive LLAP に割り当てるクラスター YARN リソースの割合 (%) と、Hive LLAP 割り当ての対象と見なすタスクノードやコアノードの数は設定できます。

たとえば、次の設定では、3 つのタスクノードまたはコアノードで 3 つのデーモンを使用して Hive LLAP を起動し、3 つのコアノードまたはタスクノードの YARN リソースの 40% を Hive LLAP デーモンに割り当てます。

```

[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true",
      "hive.llap.percent-allocation": "0.4",
      "hive.llap.num-instances": "3"
    }
  }
]

```

次の hive-site 設定を分類 API で使用し、デフォルトの LLAP リソース設定を上書きできます。

プロパティ	説明
hive.llap.daemon.yarn.container.mb	LLAP デーモンコンテナの合計サイズ (MB 単位)
hive.llap.daemon.memory.per.instance.mb	LLAP デーモンコンテナ内の実行プログラムで使用する合計メモリ (MB 単位)
hive.llap.io.memory.size	LLAP 入力/出力のキャッシュサイズ
hive.llap.daemon.num.executors	LLAP デーモンあたりの実行プログラム数

## クラスターで Hive LLAP を手動で起動する

LLAP で使用するすべての依存関係と設定は、クラスターの起動の一環として、LLAP tar アーカイブにパッケージ化されます。"hive.llap.enabled": "true" を使用して LLAP を有効にする場合は、Amazon EMR 再設定を使用して LLAP の設定を変更することをお勧めします。

それ以外の場合、hive-site.xml を手動で変更するには、hive --service llap コマンドを使用して LLAP tar アーカイブを再構築する必要があります。次に例を示します。

```
# Define how many resources you want to allocate to Hive LLAP

LLAP_INSTANCES=<how many llap daemons to run on cluster>
LLAP_SIZE=<total container size per llap daemon>
LLAP_EXECUTORS=<number of executors per daemon>
LLAP_XMX=<Memory used by executors>
LLAP_CACHE=<Max cache size for IO allocator>

yarn app -enableFastLaunch

hive --service llap \
--instances $LLAP_INSTANCES \
--size ${LLAP_SIZE}m \
--executors $LLAP_EXECUTORS \
--xmx ${LLAP_XMX}m \
--cache ${LLAP_CACHE}m \
--name llap0 \
--auxhbase=false \
--startImmediately
```

## Hive LLAP のステータスを確認する

Hive を使用して Hive LLAP のステータスを確認するには、次のコマンドを使用します。

```
hive --service llapstatus
```

YARN を使用して Hive LLAP のステータスを確認するには、次のコマンドを使用します。

```
yarn app -status (name-of-llap-service)

# example:
```



```
yarn app -status llap0 | jq
```

## Hive LLAP を起動または停止する

Hive LLAP は永続的な YARN サービスとして実行されるため、Hive LLAP を停止または再起動するには、YARN サービスを停止または再起動します。次のコマンドはこれを示しています。

```
yarn app -stop llap0  
yarn app -start llap0
```

## Hive LLAP デーモンの数を変更する

LLAP インスタンスの数を減らすには、次のコマンドを使用します。

```
yarn app -flex llap0 -component llap -1
```

詳細については、「[Flex a component of a service](#)」を参照してください。

## Hive での暗号化

このセクションでは、Amazon EMR がサポートする暗号化タイプについて説明します。

### Hive の Parquet モジュラー暗号化

Parquet モジュラー暗号化では、列レベルのアクセス制御と暗号化が可能になり、Parquet ファイル形式で保存されたデータのプライバシーとデータ整合性が強化されます。この機能は、リリース 6.6.0 以降の Amazon EMR Hive で使用できます。

ファイルの暗号化やストレージレイヤーの暗号化など、これまでサポートされていたセキュリティと整合性のソリューションについては、「Amazon EMR Management Guide」の「[Encryption Options](#)」で説明されています。これらのソリューションは Parquet ファイルにも使用できますが、統合された Parquet 暗号化メカニズムの新機能を活用することで、列レベルへのきめ細かなアクセスが可能になり、パフォーマンスとセキュリティも向上します。この機能について詳しくは、Apache GitHub のページ「[Parquet Modular Encryption](#)」を参照してください。

ユーザーは Hadoop 設定を使用して Parquet のリーダーとライターに設定を渡します。ユーザーがリーダーとライターを設定して暗号化を有効にしたり、高度な機能を切り替えたりするため

の詳細な設定については、「[PARQUET-1854: Properties-driven Interface to Parquet Encryption Management](#)」に記載されています。

## 使用例

次の例では、暗号化キーの管理に AWS KMS を使用する Hive テーブルの作成と書き込みについて説明しています。

1. ドキュメント[PARQUET-1373: 暗号化キー管理ツール](#)で説明されているように、AWS KMS のサービス `KmsClient` に `を`実装します。次の例は、実装スニペットを示しています。

```
package org.apache.parquet.crypto.keytools;

import com.amazonaws.AmazonClientException;
import com.amazonaws.AmazonServiceException;
import com.amazonaws.regions.Regions;
import com.amazonaws.services.kms.AWSKMS;
import com.amazonaws.services.kms.AWSKMSClientBuilder;
import com.amazonaws.services.kms.model.DecryptRequest;
import com.amazonaws.services.kms.model.EncryptRequest;
import com.amazonaws.util.Base64;
import org.apache.hadoop.conf.Configuration;
import org.apache.parquet.crypto.KeyAccessDeniedException;
import org.apache.parquet.crypto.ParquetCryptoRuntimeException;
import org.apache.parquet.crypto.keytools.KmsClient;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;

import java.nio.ByteBuffer;
import java.nio.charset.Charset;
import java.nio.charset.StandardCharsets;

public class AwsKmsClient implements KmsClient {

    private static final AWSKMS AWSKMS_CLIENT = AWSKMSClientBuilder
        .standard()
        .withRegion(Regions.US_WEST_2)
        .build();
    private static final Logger LOG = LoggerFactory.getLogger(AwsKmsClient.class);

    private String kmsToken;
    private Configuration hadoopConfiguration;
```

```
@Override
public void initialize(Configuration configuration, String kmsInstanceID, String
kmsInstanceURL, String accessToken) throws KeyAccessDeniedException {
    hadoopConfiguration = configuration;
    kmsToken = accessToken;
}

@Override
public String wrapKey(byte[] keyBytes, String masterKeyIdentifier) throws
KeyAccessDeniedException {
    String value = null;
    try {
        ByteBuffer plaintext = ByteBuffer.wrap(keyBytes);

        EncryptRequest req = new
EncryptRequest().withKeyId(masterKeyIdentifier).withPlaintext(plaintext);
        ByteBuffer ciphertext = AWSKMS_CLIENT.encrypt(req).getCiphertextBlob();

        byte[] base64EncodedValue = Base64.encode(ciphertext.array());
        value = new String(base64EncodedValue, Charset.forName("UTF-8"));
    } catch (AmazonClientException ae) {
        throw new KeyAccessDeniedException(ae.getMessage());
    }
    return value;
}

@Override
public byte[] unwrapKey(String wrappedKey, String masterKeyIdentifier) throws
KeyAccessDeniedException {
    byte[] arr = null;
    try {
        ByteBuffer ciphertext =
ByteBuffer.wrap(Base64.decode(wrappedKey.getBytes(StandardCharsets.UTF_8)));
        DecryptRequest request = new
DecryptRequest().withKeyId(masterKeyIdentifier).withCiphertextBlob(ciphertext);
        ByteBuffer decipheredtext =
AWSKMS_CLIENT.decrypt(request).getPlaintext();
        arr = new byte[decipheredtext.remaining()];
        decipheredtext.get(arr);
    } catch (AmazonClientException ae) {
        throw new KeyAccessDeniedException(ae.getMessage());
    }
    return arr;
}
```

```
}  
}
```

- 「デベロッパーガイド」の「キーの作成」で説明されているように、フッターの AWS KMS 暗号化キーと、IAM ロールがアクセスできる列を作成します。 <https://docs.aws.amazon.com/kms/latest/developerguide/create-keys.html> AWS Key Management Service デフォルトの IAM ロールは EMR\_ECS\_default です。
- Amazon EMR クラスター上の Hive アプリケーションで、[Apache Hive Resources ドキュメント](#)で説明されているように、ADD JAR ステートメントを使用して上記のクライアントを追加します。ステートメントの例を以下に示します。

```
ADD JAR 's3://location-to-custom-jar';
```

別の方法として、ブートストラップアクションを使用して JAR を Hive の auxlib に追加する方法があります。ブートストラップアクションに追加する行の例を以下に示します。

```
aws s3 cp 's3://location-to-custom-jar' /usr/lib/hive/auxlib
```

- 以下の設定を行います。

```
set  
  parquet.crypto.factory.class=org.apache.parquet.crypto.keytools.PropertiesDrivenCryptoFactory;  
set  
  parquet.encryption.kms.client.class=org.apache.parquet.crypto.keytools.AwsKmsClient;
```

- Parquet 形式で Hive テーブルを作成し、SERDEPROPERTIES で AWS KMS キーを指定し、それにデータを挿入します。

```
CREATE TABLE my_table(name STRING, credit_card STRING)  
ROW FORMAT SERDE 'org.apache.hadoop.hive ql.io.parquet.serde.ParquetHiveSerDe'  
WITH SERDEPROPERTIES (  
  'parquet.encryption.column.key'=<aws-kms-key-id-for-column-1>: credit_card',  
  'parquet.encryption.footer.key'='<aws-kms-key-id-for-footer>')  
STORED AS parquet  
LOCATION "s3://<bucket/<warehouse-location>/my_table";  
  
INSERT INTO my_table SELECT  
java_method ('org.apache.commons.lang.RandomStringUtils', 'randomAlphabetic', 5) as  
name,
```

```
java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',10) as
  credit_card
from (select 1) x lateral view posexplode(split(space(100),' ')) pe as i,x;

select * from my_table;
```

6. AWS KMS キーにアクセスできない (IAM ロールのアクセス拒否など) 同じ場所に外部テーブルを作成する場合、データを読み取れないことを確認します。

```
CREATE EXTERNAL TABLE ext_table (name STRING, credit_card STRING)
ROW FORMAT SERDE 'org.apache.hadoop.hive ql.io.parquet.serde.ParquetHiveSerDe'
STORED AS parquet
LOCATION "s3://<bucket>/<warehouse-location>/my_table";

SELECT * FROM ext_table;
```

7. 最後のステートメントでは次の例外が発生するはずですが。

```
Failed with exception
java.io.IOException:org.apache.parquet.crypto.KeyAccessDeniedException: Footer key:
access denied
```

## HiveServer2 での転送時の暗号化

Amazon EMR リリース 6.9.0, HiveServer2 (HS2) 以降は、[HiveServer2 での転送時の暗号化](#)セキュリティ設定の一部として TLS/SSL が有効になっています。これは、送信中の暗号化が有効になっている Amazon EMR クラスターで実行されている HS2 への接続方法に影響します。HS2 に接続するには、JDBC URL の TRUSTSTORE\_PATH および TRUSTSTORE\_PASSWORD パラメータ値を変更する必要があります。次の URL は、必須パラメータを含む HS2 の JDBC 接続の例です。

```
jdbc:hive2://HOST_NAME:10000/
default;ssl=true;sslTrustStore=TRUSTSTORE_PATH;trustStorePassword=TRUSTSTORE_PASSWORD
```

以下のクラスター上またはクラスター外の HiveServer2 暗号化には、適切なインストラクションを使用してください。

### On-cluster HS2 access

プライマリノードに SSH 接続した後に Beeline クライアントを使用して HiveServer2 にアクセスする場合は、`/etc/hadoop/conf/ssl-server.xml`を参照して、設定 TRUSTSTORE\_PATH

および `ssl.server.truststore.location` および `TRUSTSTORE_PASSWORD` パラメータ値を検索します `ssl.server.truststore.password`。

以下のコマンド例は、これらの設定を取得するのに役立ちます。

```
TRUSTSTORE_PATH=$(sed -n '/ssl.server.truststore.location/,+2p' /etc/hadoop/conf/ssl-server.xml | awk -F "[><]" '/value/{print $3}')
TRUSTSTORE_PASSWORD=$(sed -n '/ssl.server.truststore.password/,+2p' /etc/hadoop/conf/ssl-server.xml | awk -F "[><]" '/value/{print $3}')
```

## Off-cluster HS2 access

Amazon EMR クラスター外のクライアントから HiveServer2 にアクセスする場合、次のいずれかの方法を使用して `TRUSTSTORE_PATH` と `TRUSTSTORE_PASSWORD` を取得できます。

- [セキュリティ設定](#)時に作成された PEM ファイルを JKS ファイルに変換し、同じファイルを JDBC 接続 URL で使用します。例えば、`openssl` と `keytool` の場合は、次のコマンドを使用します。

```
openssl pkcs12 -export -in trustedCertificates.pem -inkey privateKey.pem -out
trustedCertificates.p12 -name "certificate"
keytool -importkeystore -srckeystore trustedCertificates.p12 -srcstoretype pkcs12
-destkeystore trustedCertificates.jks
```

- または、`/etc/hadoop/conf/ssl-server.xml` を参照して、設定 `ssl.server.truststore.location` と `ssl.server.truststore.password` を使用して `TRUSTSTORE_PATH` および `TRUSTSTORE_PASSWORD` のパラメータ値を見つけます。トラストストアファイルをクライアントマシンにダウンロードし、クライアントマシン上のパスを `TRUSTSTORE_PATH` として使用します。

Amazon EMR クラスター外のクライアントからアプリケーションにアクセスする方法の詳細については、「[Use the Hive JDBC driver](#)」を参照してください。

## Hive リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Hive のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または

[Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

## Hive バージョン情報

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.36.2	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
		worker, zookeeper-client, zookeeper-server
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-5.36.1	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.36.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.35.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server



Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.3.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.2.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Amazon EMR リリースラベル	Hive のバージョン	Hive でインストールされるコンポーネント
emr-4.1.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpsfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.0.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpsfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

## Hive リリースノート (バージョン別)

### トピック

- [Amazon EMR 7.1.0 - Hive リリースノート](#)
- [Amazon EMR 6.15.0 - Hive リリースノート](#)
- [Amazon EMR 6.14.0 - Hive リリースノート](#)
- [Amazon EMR 6.13.0 - Hive リリースノート](#)
- [Amazon EMR 6.12.0 - Hive リリースノート](#)
- [Amazon EMR 6.11.0 - Hive リリースノート](#)

- [Amazon EMR 6.10.0 - Hive リリースノート](#)
- [Amazon EMR 6.9.0 - Hive リリースノート](#)
- [Amazon EMR 6.8.0 - Hive リリースノート](#)
- [Amazon EMR 6.7.0 - Hive リリースノート](#)
- [Amazon EMR 6.6.0 - Hive リリースノート](#)

## Amazon EMR 7.1.0 - Hive リリースノート

### Amazon EMR 7.1.0 - Hive の変更点

型	説明
バグ修正	<a href="#">HIVE-24381</a> – スキップヘッダー/フッターが含まれている場合、圧縮テキスト入力は 0 行を返します。
バグ修正	<a href="#">HIVE-24190</a> – LLAP: ShuffleHandler TEZ-4233 に従って DISK_ERROR_EXCEPTION を返す可能性があります。
バグ修正	<a href="#">HIVE-23073</a> – シェード・ネティ。
バグ修正	<a href="#">HIVE-23073</a> – Netty をシェードし、netty 4.1.48.Final にアップグレードします。
バグ修正	<a href="#">HIVE-23148</a> – ネットティシェーディングにより、外部クライアントフローが切断されています。
バグ修正	<a href="#">HIVE-25180</a> – Netty をアップグレードします。
バグ修正	<a href="#">HIVE-24524</a> – LLAP ShuffleHandler: Netty4 にアップグレードし、可能な場合は Netty3 の依存関係を hive から削除します。
バグ修正	<a href="#">HIVE-28000</a> – Hive QL: 型強制を実行できない場合、「not in」句は誤った結果を提供します。

型	説明
バグ修正	<a href="#">HIVE-27993</a> – Netty4 ShuffleHandler は 1 つのボススレッドを使用する必要があります。
Upgrade	Netty を 4.1.100.Final にアップグレード
Upgrade	Jetty を 9.4.53.v20231009 にアップグレード
Upgrade	Zookeeper を 3.9.1 にアップグレード

### Amazon EMR 7.1.0 - Hive の変更点

- Amazon EMR 7.1 は、Hive を Netty 4.1.100.Final Netty3 のセキュリティの脆弱性を解決します。hive-druid-handler は netty3 に依存するため、Hive には Amazon EMR 7.1 の Hive のクラスパスに hive-druid-handler JAR がありません。Druid ハンドラーが Netty の 4.1.100.Final 以降のバージョンをサポートすると、今後の Amazon EMR リリースで Hive のクラスパスに含まれます。Amazon EMR リリース 7.1 以降で hive-druid-handler JAR AWS が必要な場合は、サポートにお問い合わせください。

### Amazon EMR 6.15.0 - Hive リリースノート

#### Amazon EMR 6.15.0 - Hive の変更点

型	説明
機能	<a href="#">TEZ-4397</a> のサポート - Tez 非同期スプリットオープンの場合、Hive は <a href="#">Tez 非同期分割オープニング</a> で説明されている Tez 設定をサポートするようになりました。
バグ修正	<a href="#">HIVE-25400</a> - BytesColumnVector のオフセット更新を setValPreallocated に移動します。
バグ修正	<a href="#">HIVE-25190</a> - BytesColumnVector での多数の小さな割り当てを修正しました。

型	説明
Upgrade	Apache Hadoop を 3.3.6 に更新します。
Upgrade	<a href="#">HIVE-26684</a> – maven-shade-plugin を 3.4.1 へアップグレードします。
改良点	Amazon EMR クラスターの起動時間を短縮するには、HCatalog スタートアップスクリプトから 15 秒のスリープ時間を削除します。

## Amazon EMR 6.14.0 - Hive リリースノート

### Amazon EMR 6.14.0 - Hive の変更点

型	説明
改良点	<a href="#">HIVE-26762</a> : HiveFilterSetOpTransposeRule のオペランドのプルーニングを削除
バグ修正	<a href="#">HIVE-27582</a> : で HBase テーブル入力形式をキャッシュしない FetchOperator
バグ修正	<a href="#">HIVE-26452</a> : JOIN を MAPJOIN に変換し、JOIN 列が複数回参照される場合の NPE
バグ修正	<a href="#">HIVE-26416</a> : AcidUtils.isRawFormatFile() が non-ORC ファイルに対して InvalidProtocolBufferException をスローする
バグ修正	<a href="#">HIVE-26105</a> : [コメント] 列に特定の漢字が含まれている場合、[列の表示] で追加の値が表示される

型	説明
バグ修正	<a href="#">HIVE-25864</a> : Hive クエリの最適化により、ウィンドウ関数による述語プッシュダウンの間違ったプランが生成される
バグ修正	<a href="#">HIVE-25224</a> : 異なる bucketing_versions があるテーブルを含む複数の INSERT ステートメントがエラーになる
バグ修正	<a href="#">HIVE-24151</a> : 文字列に non-ASCII 文字が含まれていると MultiDelimitSerDe がデータをシフトする
バグ修正	<a href="#">HIVE-23606</a> : (LLAP) EncodedReaderImpl の DirectByteBuffer クリーンアップの遅延
バグ修正	<a href="#">HIVE-22165</a> : <a href="#">HIVE-14296</a> によって <code>SessionManager.closeSession</code> に導入された同期により、ビジー状態の Hive サーバーで高いレイテンシーが発生する
バグ修正	<a href="#">HIVE-21304</a> : バケット化バージョンの使用をより堅牢にする

## Amazon EMR 6.13.0 - Hive リリースノート

### Amazon EMR 6.13.0 - Hive の変更点

型	説明
改良点	Python スクリプトを Python3 をサポートするようにアップグレード
改良点	<a href="#">HIVE-27097</a> : MetaStore クライアントとサーバーの再試行戦略を改善



型	説明
バグ修正	<a href="#">HIVE-21778</a> : CBO: 「構造体が null ではない」が Null を許容として評価され、クエリで常にフィルターミスが生じる
バグ修正	<a href="#">HIVE-21009</a> : ユーザーがバインドユーザーを設定する機能の追加
バグ修正	<a href="#">HIVE-22661</a> : データが path に読み込まれたバケット化されていないテーブルで圧縮が失敗する
バグ修正	<a href="#">HIVE-19718</a> : パーティションを一括追加すると、各パーティションのテーブルも取得される
バグ修正	<a href="#">HIVE-22173</a> : 複数の側面ビューを含むクエリがコンパイル中にハングする
バグ修正	<a href="#">HIVE-27088</a> : ポストジョインフィルター付きの内部ジョインと外部ジョインをマージすると結果が不正確になる
バグ修正	<a href="#">HIVE-21935</a> : Hive のベクトル化: UDF のベクトル化によるパフォーマンスの低下
バグ修正	<a href="#">HIVE-25299</a> : UTC 以外のタイムゾーンで、タイムスタンプを数値データ型にキャストすると正しくない
バグ修正	<a href="#">HIVE-24626</a> : LLAP: すべての IO エレベータースレッドがビジー状態で、キューがいっぱいの別のリーダーへのキュー登録を行うと、リーダースレッドが枯渇する可能性がある
バグ修正	<a href="#">HIVE-27029</a> : Hive クエリがファイルシステムのクローズエラーで失敗し、HIVE-26352 の再作業が行われた

型	説明
バグ修正	<a href="#">HIVE-26352</a> : 圧縮時に GSS 例外が発生して Tez キューのアクセスチェックが失敗する
バグ修正	<a href="#">HIVE-24590</a> : 操作ログからまだ log4j アペンダーがリークする
バグ修正	<a href="#">HIVE-24552</a> : で発生する可能性のある HMS 接続のリークまたは蓄積 loadDynamicPartitions
バグ修正	<a href="#">HIVE-27069</a> : バケットマップ結合の結果が正しくない
バグ修正	<a href="#">HIVE-27344</a> : RecordReaderImpl#close に null チェックを追加する
バグ修正	<a href="#">HIVE-27439</a> : 10 進数でのスペースのサポート
バグ修正	<a href="#">HIVE-27267</a> : サブクエリを使用して 10 進数のバケット列でバケットマップ結合を実行すると、結果が不正になる
バグ修正	<a href="#">HIVE-21986</a> : HiveServer Web UI: デフォルトのレスポンスヘッダーで Strict-Transport-Security を設定する
バグ修正	<a href="#">HIVE-22148</a> : S3A 委任トークンがコンパクターのジョブ設定に追加されない。
バグ修正	<a href="#">HIVE-22622</a> : Hive では、重複する属性名を持つ構造体を作成できる
バグ修正	<a href="#">HIVE-22008</a> : LIKE 演算子は複数行の入力と一致する必要がある
バグ修正	<a href="#">HIVE-23144</a> : LLAP: サービス停止時の QueryTracker クリーンアップを許可する serviceStop

型	説明
バグ修正	<a href="#">HIVE-22391</a> : Hive のクエリ結果キャッシュを チェック中の NPE
バグ修正	競合状態による <a href="#">HIVE-23305</a> : NullPoint erException in LlapTaskSchedulerService addNode
バグ修正	<a href="#">HIVE-22178</a> : Parquet が CastException の後に FilterPredicate スロー SchemaEvolution
バグ修正	<a href="#">HIVE-21517</a> : 修正 AggregateStatsCache
バグ修正	<a href="#">HIVE-21825</a> : アクティブ/パッシブ HA が有効に なっている場合のクライアントエラーメッセ ージを改善
バグ修正	<a href="#">HIVE-23389</a> : FilterMergeRule は AssertionError
バグ修正	<a href="#">HIVE-22767</a> : Beeline がコメント内のセミコロ ンを正しく解析しない
バグ修正	<a href="#">HIVE-22996</a> : BasicStats parsing は NULL ま たは空の文字列をプロアクティブにチェックす る必要があります
バグ修正	<a href="#">HIVE-22808</a> : HiveRelFieldTrimmer 処理しない HiveTableFunctionScan
バグ修正	<a href="#">HIVE-22437</a> : メタデータのロックに関する LLAP メタデータキャッシュ NPE。
バグ修正	INFO レベルでの <a href="#">HIVE-22606</a> : AvroSerde logs avro.schema.literal
バグ修正	<a href="#">HIVE-22713</a> : Join-Fil(*)-RS 構造体では継続的 な伝播を行うべきではない

型	説明
バグ修正	<a href="#">HIVE-21624</a> : LLAP: スレッドレベルの CPU メトリクスが壊れている
バグ修正	<a href="#">HIVE-22815</a> : MROutput での不要なファイルシステムオブジェクトの作成を減少
バグ修正	<a href="#">HIVE-23060</a> : 次のエラーでクエリが失敗する。 「Grouping sets expression is not in GROUP BY key. Error encountered near token」
バグ修正	<a href="#">HIVE-22236</a> : NOT IN サブクエリを含むビューを選択するビューを作成できない
バグ修正	<a href="#">HIVE-19886</a> : —hiveconf hive.log.file を使用すると、ログが 2 つのファイルに送信されることがある
バグ修正	<a href="#">HIVE-20620</a> : 動的パーティショニングを使用してバケット化されたソート済み MM テーブルに挿入すると、マニフェストの衝突が発生する
バグ修正	<a href="#">HIVE-14557</a> : と Mapjoin の両方 SkewJoin が有効になっている場合の Nullpointer
バグ修正	<a href="#">HIVE-20471</a> : デフォルトのデータベースパスを取得する際に問題が発生する
バグ修正	<a href="#">HIVE-20598</a> : 計算の HiveAlgorithmsUtil 誤字を修正
バグ修正	<a href="#">HIVE-14737</a> : Kerberized Hive サーバー 2 ウェブ UI の /logs へのアクセスに関する問題
バグ修正	<a href="#">HIVE-20733</a> : GenericUDFOPEqualNS はプランの説明で = を使用しない場合がある

型	説明
バグ修正	<a href="#">HIVE-20848</a> : 設定後 UpdateInputAccessTimeHook、クエリが失敗し、テーブルが見つかりません。
バグ修正	<a href="#">HIVE-18929</a> : HiveStringUtils.java humanReadableInt のメソッドには競合状態があります。
バグ修正	<a href="#">HIVE-20841</a> : LLAP: 動的ポートを設定可能にする
バグ修正	<a href="#">HIVE-20930</a> : VectorCoalesce FILTER モードは有効になりません
バグ修正	<a href="#">HIVE-21007</a> : セミ結合 + ユニオンは間違ったプランにつながる可能性がある
バグ修正	<a href="#">HIVE-21074</a> : IS NOT NULL 条件では Hive バケットテーブルのクエリプルーニングが機能しない
バグ修正	<a href="#">HIVE-21223</a> : パーティションが存在しない場合、null パーティション CachedStore を返します
バグ修正	<a href="#">HIVE-19625</a> : NPE が発生する可能性があり、Hive#copyFiles に実際の例外が隠れている
バグ修正	<a href="#">HIVE-17020</a> : 重度 RS 重複排除により OP ツリーのブランチが誤って削除されることがある
バグ修正	<a href="#">HIVE-20168</a> : ReduceSinkOperator ログ記録が非表示
バグ修正	<a href="#">HIVE-20879</a> : 射影式で null を使用すると、CastException

型	説明
バグ修正	<a href="#">HIVE-20888</a> : TxnHandler: イミュータブルリストで呼び出された sort()
バグ修正	<a href="#">HIVE-19948</a> : HiveCli 引用符が文字列内にある場合、コマンドをセミコロンで適切に分割していない
バグ修正	<a href="#">HIVE-20621</a> : resultset.next で GetOperationStatus 呼び出され、段階的な速度低下を引き起こす
バグ修正	<a href="#">HIVE-20854</a> : 妥当なデフォルト: Hive の Zookeeper ハートビート間隔は 20 分だが、2 に変更する
バグ修正	<a href="#">HIVE-20330</a> : H は複数の入力を持つジョブの複数の InputJobInfo オブジェクトを処理 CatLoader できない
バグ修正	<a href="#">HIVE-20787</a> : MapJoinBytesTableContainer dummyRow ケースは再利用を処理しない
バグ修正	<a href="#">HIVE-20331</a> : すべてユニオン、側面ビュー、および結合を使用したクエリが「cannot find parent in the child operator」で失敗する
バグ修正	<a href="#">HIVE-19968</a> : UDF 例外がスローされない
バグ修正	<a href="#">HIVE-20410</a> : トランザクションテーブルでの Insert Overwrite が中止されると、「Not enough history available for...」というエラーが発生する
バグ修正	<a href="#">HIVE-20059</a> : Hive ストリーミングでは、例外発生時に無条件にシェードプレフィックスを試す必要がある

型	説明
バグ修正	<a href="#">HIVE-19424</a> : NPE In MetaDataFormatters
バグ修正	<a href="#">HIVE-20355</a> : HiveConnection.setSchema のクリーンアップパラメータ
バグ修正	<a href="#">HIVE-20858</a> : ユーティリティの設定でシリアライザーが正しく初期化されていません。createEmptyBuckets
バグ修正	<a href="#">HIVE-20424</a> : schematool は beeline 履歴を汚染してはならない
バグ修正	<a href="#">HIVE-20338</a> : LLAP: POSIX ミューテーションセマンティクスを使用した HDFS プロトコル impls があるファイルシステムに対して合成ファイル ID を強制する
バグ修正	<a href="#">HIVE-11708</a> : 論理演算子が NULL ClassCast Exceptions で増加
バグ修正	<a href="#">HIVE-21082</a> : HPL/SQL では、宣言ステートメントは文字型の変数をサポートしていない
バグ修正	<a href="#">HIVE-16690</a> : LLAP クラスターサイズに基づいて Tez 直積集合エッジを設定する
バグ修正	<a href="#">HIVE-21296</a> : varchar パーティションをドロップすると例外が発生する
バグ修正	<a href="#">HIVE-14516</a> : OrclInputFormat.SplitGenerator.callInternal
バグ修正	<a href="#">HIVE-20981</a> : ストリーミング/AbstractRecordWriterリーク HeapMemoryMonitor
バグ修正	<a href="#">HIVE-20043</a> : HiveServer2: SessionState の周囲に静的同期ブロックがある AtomicBoolean

型	説明
バグ修正	パッチが空の場合、 <a href="#">HIVE-20191</a> : PreCommit パッチアプリケーションは失敗しない
バグ修正	<a href="#">HIVE-20400</a> : FS があいまいになる可能性を避けるため、テーブル作成では常に完全修飾パスを使用する必要がある
バグ修正	スキューされた列にアクセスする前に skewedInfo の NULL チェックを追加

## Amazon EMR 6.12.0 - Hive リリースノート

### Amazon EMR 6.12.0 - Hive の変更点

型	説明
改良点	JDK 11 と JDK 17 ランタイムのサポートが追加されました
改良点	S3 Select を使用する際に、大文字と小文字が区別されるキーワード列名と予約キーワード列名をクエリするサポートが追加されました。これを使用するには、"s3select.column.mapping" = " <i>column1:fieldName1</i> , <i>column2:fieldName2</i> , ..." という形式でテーブルプロパティを定義してください。
改良点	<a href="#">HIVE-23133</a> : 数値演算はハードウェアアーキテクチャによって結果が異なる場合がある
改良点	<a href="#">HIVE-27145</a> : HIVE-23133 のフォローアップとして残りの数学関数 StrictMath に使用
バグ修正	EMR Hive 6.4.0 で <a href="#">HIVE-22900</a> を移植したために発生した get_partitions_by_filter と



型	説明
	get_num_partitions_by_filter HMS API のワイルドカードの非互換性を修正
バグ修正	<a href="#">HIVE-26736</a> : WITH 句を含むネストされたビューの認証が失敗する
バグ修正	<a href="#">HIVE-22416</a> : parallel 実行が有効になっていると、MR 関連の操作ログが見つからない
バグ修正	<a href="#">HIVE-19653</a> : グループ化セットを使用した groupby の述語プッシュダウンが正しくない
バグ修正	<a href="#">HIVE-22094</a> : ClassCastException: hive ql.exec.vector.DecimalColumnVector で失敗したクエリは、hive ql.exec.vector.Decimal64 にキャストできませんColumnVector
バグ修正	<a href="#">HIVE-26340</a> : クエリに大文字のウィンドウ関数が含まれていると、ベクトル化された PTF 演算子が失敗する
バグ修正	<a href="#">HIVE-26184</a> : 一部のキーが大きく歪んでいると、GROUP BY を使用した COLLECT_SET が非常に遅くなる
バグ修正	<a href="#">HIVE-26373</a> : Avro データを使用して HBase テーブルからタイムスタンプを読み取る ClassCastException 場合
バグ修正	<a href="#">HIVE-26388</a> : CTAS クエリのソーステーブルに文字列型以外の列 ClassCastException がある場合 <a href="#">HIVE-26172</a> のアップグレード: Hive - CVE-2021-36373 および CVE-2021-36374 により Ant を 1.10.11 にアップグレード

型	説明
バグ修正	<a href="#">HIVE-26114</a> : プレフィックススペースを指定した dfs コマンドを使用すると例外が発生する jdbc 接続 hiveserver2 を修正
バグ修正	<a href="#">HIVE-26396</a> : trunc 関数の精度インターセプトに問題があり、結果に 0 が多数返される
バグ修正	<a href="#">HIVE-26446</a> : HiveProtoLoggingHook パーティションテーブルの TablesWritten フィールドへの入力に失敗しました。
バグ修正	<a href="#">HIVE-26639</a> : ConstantVectorExpression デフォルトの文字セットに依存し ExplainTask ない
バグ修正	<a href="#">HIVE-22670</a> : Parquet ファイルの読み取りにベクトル化されたリーダーが使用される ArrayIndexOutOfBoundsException 場合
バグ修正	<a href="#">HIVE-23607</a> : 権限の問題: 別のビューでのビューの作成は成功するが、ビューの変更は失敗する
バグ修正	<a href="#">HIVE-25498</a> : 個数が 31 個を超える個別関数を含むクエリが間違った結果を返す
バグ修正	<a href="#">HIVE-25780</a> : 64 を超えるグループ化セット II DistinctExpansion を作成
バグ修正	<a href="#">HIVE-23868</a> : ウィンドウ関数の仕様: 先行/後続 0 件をサポート
バグ修正	<a href="#">HIVE-24539</a> : OrclnputFormat スキーマの生成では列区切り文字を優先する必要があります
バグ修正	<a href="#">HIVE-23476</a> : LLAP: mmap の場合にもアリーナを事前に割り当てておく

型	説明
バグ修正	<a href="#">HIVE-25806</a> : LlapCacheAwareFs Parquet、L LAP IO で発生する可能性のあるリーク
バグ修正	<a href="#">HIVE-23498</a> : で HTTP トレースメソッドを無効にする ThriftHttpCliService
バグ修正	<a href="#">HIVE-25729</a> : 完全に開始されると通知 ThriftUnionObjectInspector される
バグ修正	<a href="#">HIVE-23846</a> : ビットベクトルの不必要なシリアル化と逆シリアル化を回避
バグ修正	<a href="#">HIVE-24233</a> : cbo を無効にした状態でサブクエリが NULL ポインターを投げる場合を除く
バグ修正	<a href="#">HIVE-24276</a> : HiveServer2 loggerconf.jsp Cross-Site Scripting (XSS) の脆弱性
バグ修正	<a href="#">HIVE-25721</a> : 外部結合の結果が間違っている
バグ修正	<a href="#">HIVE-25223</a> : 制限付き選択では非ネイティブテーブルの行が返されない
バグ修正	<a href="#">HIVE-25794</a> : CombineHiveRecordReader: ループ内のログステートメントがメモリ負荷につながる
バグ修正	<a href="#">HIVE-23602</a> : オペレーションハンドルセットに Java コンカレントパッケージを使用する
バグ修正	<a href="#">HIVE-24045</a> : デフォルトデータベースの作成時期に関するログがない
バグ修正	<a href="#">HIVE-24305</a> : 値が引用符で囲まれていると、avro 10 進スキーマがスケール/精度を正しく入力しない

型	説明
バグ修正	<a href="#">HIVE-25844</a> : 例外逆シリアル化エラーにより、beeline がすぐに終了することがある
バグ修正	<a href="#">HIVE-25040</a> : データベースカスケードをドロップすると永続関数を削除できない
バグ修正	<a href="#">HIVE-23501</a> : 複合型がプリミティブ型に変換 VectorDeserializeRow された場合の の AOOB
バグ修正	<a href="#">HIVE-23704</a> : Thrift HTTP サーバーが認証ハンドルを正しく処理しない
バグ修正	<a href="#">HIVE-23529</a> : row_deserialize を実行すると、uniontype の CTAS が壊れる
バグ修正	<a href="#">HIVE-24144</a> : の getIdentifierQuote文字列が不正な値 HiveDatabaseMetaData を返します
バグ修正	<a href="#">HIVE-23850</a> : サブジェクトがグループ化セットが存在する列でない場合に PPD を許可する
バグ修正	<a href="#">HIVE-24036</a> : getSplits UDF 呼び出しのプランをシリアル化中に Kryo 例外が発生した
バグ修正	<a href="#">HIVE-25919</a> : H でブール列述語をプッシュ ClassCastException する場合BaseStorageHandler
バグ修正	<a href="#">HIVE-25261</a> : RetryingHMSHandler は、ターゲットの簡単な説明 MetaException で をラップする必要があります
バグ修正	<a href="#">HIVE-24792</a> : オペレーション中にスレッドリークが発生する可能性がある

型	説明
バグ修正	<a href="#">HIVE-23409</a> : Timeline サービスダウンで TezSession アプリケーションの再オープンが失敗した場合、再試行後、TezSession からのデフォルト SessionPool は閉じられます
バグ修正	<a href="#">HIVE-23615</a> : Beeline コマンドクラスの null ポインターを参照しないでください
バグ修正	<a href="#">HIVE-24849</a> : ロケーションに大量のファイルがある場合に外部テーブルソケットのタイムアウトを作成する (3.1.2 に影響)
バグ修正	<a href="#">HIVE-24193</a> : 名前が変更された hive acid テーブルでの選択クエリでは何も出力されない
バグ修正	<a href="#">HIVE-25209</a> : SUM 関数を使用した SELECT クエリで予期しない結果が生じる
バグ修正	Groupby 演算子にグループ化セットがない場合、 <a href="#">HIVE-23666</a> : checkHashModeEfficiency はスキップされます
バグ修正	<a href="#">HIVE-23873</a> : CBO がオフの場合、Hive JDBCStorageHandler テーブルのクエリが NPE で失敗する
バグ修正	<a href="#">HIVE-24149</a> : HMS HiveStreamingConnection 接続を閉じない
バグ修正	<a href="#">HIVE-25561</a> : 強制終了されたタスクはファイルをコミットしてはならない (2.x および 3.x バージョンに影響)。
バグ修正	<a href="#">HIVE-25683</a> : AcidUtils.isRawFormatFile でリーダーを閉じる

型	説明
バグ修正	<a href="#">HIVE-24294</a> : TezSessionPool セッションがスローできる AssertionError
バグ修正	<a href="#">HIVE-24182</a> : 永続的な UDF に関する Ranger の認証の問題
バグ修正	<a href="#">HIVE-22805</a> : 条件付き配列またはマップによるベクトル化が実装されておらず、エラーが発生する
バグ修正	<a href="#">HIVE-22828</a> : Decimal64: NVL ステートメントと CASE ステートメントは暗黙的に decimal64 を 128 に変換する
バグ修正	<a href="#">HIVE-21398</a> : 推定統計情報を含む列は一意的なキーと見なすべきではない
バグ修正	<a href="#">HIVE-22490</a> : パスに特殊文字を含む jar を追加するとエラーが発生する
バグ修正	<a href="#">HIVE-22700</a> : 不正な圧縮を行うとメモリリークが発生することがある
バグ修正	<a href="#">HIVE-22053</a> : 関数作成時に関数名が正規化されない
バグ修正	<a href="#">HIVE-22595</a> : 外部スキーマのある Avro テーブルテーブルで動的パーティションの挿入が失敗する
バグ修正	<a href="#">HIVE-21795</a> : 分割テーブルで mapjoin を行うと、ローラップサマリー行が欠落することがある
バグ修正	DataTypePhysicalVariation が null VectorCoalesce の場合の <a href="#">HIVE-22987</a> : ClassCastException in

型	説明
バグ修正	ベクトル化 ArrayIndexOutOfBounds における <a href="#">HIVE-22814</a> : getDataPhysicalVariation
バグ修正	<a href="#">HIVE-22523</a> : キューがいっぱいになると、のエラーハンドラーがブロック LlapRecordReader する可能性があります
バグ修正	<a href="#">HIVE-21796</a> : ArrayWritableObjectInspector.equals には $O(2^{\text{nesting\_depth}})$ の時間がかかる場合があります
バグ修正	<a href="#">HIVE-22929</a> : パフォーマンス: 引用符で囲まれた識別子の解析では、String.replaceAll() を介して使い捨て正規表現が使用される
バグ修正	<a href="#">HIVE-21641</a> : Llap 外部クライアントが、beeline とは異なる精度/スケールの 10 進数列を返す
バグ修正	<a href="#">HIVE-22207</a> : クラスターの「dfs.block.size」が「128m NumberFormatException」の場合の Tez: SplitGenerator throws
バグ修正	<a href="#">HIVE-22114</a> : すべてのバケットが空の場合、パーティション挿入専用テーブルの挿入クエリが失敗する
バグ修正	<a href="#">HIVE-22165</a> : . SessionManagercloseSession で HIVE-14296 によって導入された同期は、ビジー状態のハイブサーバーで高いレイテンシーを引き起こします
バグ修正	<a href="#">HIVE-22744</a> : TezTask 複数のアウトエッジを持つ頂点には、比例ソートメモリが必要です
バグ修正	<a href="#">HIVE-22072</a> : テーブルを変更して列を変更しても、制約参照は更新されない

型	説明
バグ修正	<a href="#">HIVE-22075</a> : HIVE-14200 から発生していた max-reducers=1 リグレッションを修正
バグ修正	<a href="#">HIVE-22527</a> : Hive on Tez: 小さなファイルをマージするジョブが別のキュー (デフォルトキュー) に送信される
バグ修正	<a href="#">HIVE-22816</a> : QueryCache: ビューを使用するクエリは、CTE 拡張後にキャッシュされる可能性があります
バグ修正	<a href="#">HIVE-22733</a> : hive の操作ログプロパティを無効にした後も、HS2 が操作ログを保存する
バグ修正	<a href="#">HIVE-22699</a> : マスク UDF は数値 0 をマスクする必要がある
バグ修正	<a href="#">HIVE-23356</a> : グループ化セット式を含むクエリを処理している間は、ハッシュアグリゲーションが常に無効になる。
バグ修正	<a href="#">HIVE-21568</a> : HiveRelOptUtilisRowFiltering計画はプロジェクトをスキップする必要があります
バグ修正	<a href="#">HIVE-21760</a> : SMB 結合では共有ワークの最適化はバイパスする必要がある
バグ修正	<a href="#">HIVE-22712</a> : ReExec Driver execute ユーザー定義キューに関係なく、デフォルトキューでクエリを送信します
バグ修正	<a href="#">HIVE-21397</a> : BloomFilter for hive Managed [ACID] テーブルが期待どおりに機能しない



型	説明
バグ修正	<a href="#">HIVE-23011</a> : 共有ワークオブティマイザーは、結合を比較する際に残余述語をチェックする必要がある
バグ修正	<a href="#">HIVE-21412</a> : ACID PostExecOrcFileDump テーブルでは機能しない
バグ修正	<a href="#">HIVE-22201</a> : 大きなテーブルが選択され ArrayIndexOutOfBoundsException ではない場合は ConvertJoinMapJoin#checkShuffleSizeForLargeTable throws
バグ修正	<a href="#">HIVE-21971</a> : 一時的な関数 + GenericUDF を持つ ReflectionUtils`::CONSTRUCTOR_CACHE` が原因で HS2 がクラスローダーをリークする GenericUDF
バグ修正	<a href="#">HIVE-21368</a> : ベクトル化: 不必要な Decimal64 -> HiveDecimal 変換
バグ修正	<a href="#">HIVE-25416</a> : datanucleus-api-jdo バグによる Hive メタストアのメモリリーク
バグ修正	<a href="#">HIVE-22219</a> : ノードマネージャーをダウンさせると LLAP サービスの再起動がブロックされる
バグ修正	<a href="#">HIVE-21793</a> : hive.stats.fetch.column.stats が false に設定されていても CBO が列の統計情報を取得する
バグ修正	<a href="#">HIVE-22163</a> : CBO: CBO を有効にすると、推定が無効になっていても統計推定が有効になる
バグ修正	<a href="#">HIVE-18735</a> : トランザクション属性が失われるようなテーブルを作成する

型	説明
バグ修正	<a href="#">HIVE-22433</a> : Hive JDBC ストレージハンドラー: JDBC データソースから BOOLEAN と TIMESTAMP DataType から取得された結果が間違っている
バグ修正	<a href="#">HIVE-19430</a> : ObjectStorecleanNotificationEvents OutOfMemory 多数の保留中のイベント
バグ修正	<a href="#">HIVE-20785</a> : JDBC . DatabaseMetaDatagetPrimaryKeys method のキー名が正しくない
バグ修正	<a href="#">HIVE-16116</a> : beeline.properties で beeline.hiveconfvariables={} の場合、Beeline が NPE をスローする
バグ修正	<a href="#">HIVE-20066</a> : hive.load.data.owner がフルプリンシパルと比較される
バグ修正	<a href="#">HIVE-20489</a> : クエリがハングするプランを説明する
バグ修正	<a href="#">HIVE-21033</a> : オペレーションを閉じるのを忘れると、さらに HiveServer2 つの出力がカットされます
バグ修正	<a href="#">HIVE-19888</a> : からの誤解を招く「METASTORE_FILTER_HOOK will be ignored」警告 SessionState
バグ修正	<a href="#">HIVE-20303</a> : INSERT OVERWRITE TABLE db.table PARTITION (...) IF NOT EXISTS スロー InvalidTableException
バグ修正	<a href="#">HIVE-16144</a> : CompactionInfo equals/hashCode はありませんが、Set で使用されています

型	説明
バグ修正	<a href="#">HIVE-20818</a> : WHERE サブクエリで作成されたビューは、サブクエリで参照されるビューを直接入力と見なす
バグ修正	<a href="#">HIVE-21005</a> : LLAP: 分割ごとのより多くのストライプリークの読み取り ZlibCodecs
バグ修正	<a href="#">HIVE-20771</a> : 空の構造体で LazyBinarySerDe 失敗します。
バグ修正	<a href="#">HIVE-18852</a> : 変更テーブルの検証で誤解を招くエラーメッセージ
バグ修正	<a href="#">HIVE-21124</a> : HPL/SQL は CREATE TABLE LIKE ステートメントをサポートしない
バグ修正	<a href="#">HIVE-20935</a> : llap パッケージ tarball のアップロードが EC2 で失敗し、LLAP サービスの起動に失敗する
バグ修正	<a href="#">HIVE-20409</a> : Hive ACID: 更新/削除/マージを行っても hdfs ステージングディレクトリがクリーンアップされない
バグ修正	<a href="#">HIVE-20570</a> : hive.optimize.union.remove=true ですべてをユニオンすると、プランが正しくない
バグ修正	<a href="#">HIVE-20421</a> : hive-default.xml.template 内の文字エンティティ '\b' が不正
バグ修正	<a href="#">HIVE-19133</a> : HS2 WebUI のフェーズ単位のパフォーマンスメトリックが正しく表示されない
バグ修正	<a href="#">HIVE-18977</a> : JDO とダイレクト SQL では、パーティションを一覧表示すると異なる結果が返される

型	説明
バグ修正	<a href="#">HIVE-20034</a> : 下位互換性のための MetaStore 例外処理の変更をロールバックする
バグ修正	<a href="#">HIVE-20672</a> : のログ記録スレッドは、固定間隔ごとに報告 LlapTaskSchedulerService する必要があります
バグ修正	<a href="#">HIVE-12812</a> : 集約関数によるユニオンをサポートするため、mapred.input.dir.recursive をデフォルトで有効にする
バグ修正	<a href="#">HIVE-20147</a> : Hive ストリーミング取り込みは同期ログに基づいて行われる
バグ修正	<a href="#">HIVE-19203</a> : におけるスレッド安全性の問題 HiveMetaStore
バグ修正	<a href="#">HIVE-20091</a> : Tez: FileSinkOperator 出力用のセキュリティ認証情報を追加する
バグ修正	<a href="#">HIVE-16906</a> : Hive ATSHook は ATS に接続する前に yarn.timeline-service.enabled をチェックする必要がある
バグ修正	<a href="#">HIVE-20714</a> : 1 つのプロパティの SHOW tblproperties を行うと、名前列の値が返される
バグ修正	<a href="#">HIVE-24730</a> : Shims クラスは hive-site.xml と tez-site.xml の値をサイレントにオーバーライドする
バグ修正	<a href="#">HIVE-22055</a> : テキストファイルからデータをロードした後、選択カウントが不適切な結果になる

## Amazon EMR 6.11.0 - Hive リリースノート

### Amazon EMR 6.11.0 - Hive の変更点

型	説明
改良点	パーティションの削除のパフォーマンスを向上させるため、マルチスレッドによるパーティションの削除のサポートを追加
改良点	エンコードされた Hive クエリファイルの読み取りをサポート
改良点	Hive on Tez ジョブの Tez シャッフルハンドラーをデフォルトで有効化
バグ	hive.groupby.skewindata が有効になっているときの誤った結果を修正するために、リデューサーへのキーの確定的配布を有効にするオプションを追加 ( <a href="#">HIVE-20220</a> で報告)
バグ	デフォルトのパーティション名が設定されている場合の統計情報の計算が失敗する問題を修正
バグ	転送時の暗号化が有効になっているクラスターで、SSL が HiveServer2 用にすぐに設定されるときに渡されるカスタム SSL 分類パラメータをすべて尊重する
バックポート	<a href="#">HIVE-23617</a> : storage-api FindBug の問題を修正
バックポート	<a href="#">HIVE-26408</a> : ベクトル化: スクラッチ列の割り当て解除を修正し、子を出力 ConstantVectorExpression として再利用しない
バックポート	<a href="#">HIVE-23614</a> : 常に HiveConfig に渡す removeTempOrDuplicateFiles

型	説明
バックポート	<a href="#">HIVE-23354</a> : からファイルサイズのサニティチェックを削除する compareTempOrDuplicateFiles
バックポート	<a href="#">HIVE-20344</a> : をスローする SBA Privilege Synchronizer を修正しました AccessControlException。また、特権シンクロナイザーを無効にするプロパティ hive.privilege.synchronizer も導入
バックポート	<a href="#">HIVE-15826</a> : すべてのの「serialization.encoding」の設定をサポート SerDes
バックポート	<a href="#">HIVE-18284</a> : dynpart ソート最適化で「distribute by」句を使用してデータを挿入するときの NPE を修正
バックポート	<a href="#">HIVE-24930</a> : 子演算からの Operator.setDone() ショートサーキットはベクトル化されたコードパスでは使用されない (childSize == 1 の場合)
バックポート	<a href="#">HIVE-24523</a> : のベクトル化された読み取りパス LazySimpleSerde は、タイムスタンプの SERDEPROPERTIES を尊重しません
バックポート	<a href="#">HIVE-23265</a> : 制限とオフセットが設定された重複行セットが返される
バックポート	<a href="#">HIVE-21492</a> : thrift/custom ツールを使用して生成された Parquet ファイルを読み取れ VectorizedParquetRecordReader ない
バックポート	<a href="#">HIVE-22540</a> : ベクトル化: Decimal64 列は VectorizedBatchUtil.makeLikeColumnVector() では機能しません

型	説明
バックポート	<a href="#">HIVE-22588</a> : ベクトルグループ化モードを切り替えたときに、残りのグループ化セットの残りの行をフラッシュする
バックポート	<a href="#">HIVE-22551</a> : BytesColumnVector initBuffer はベクトルと長さを一貫してクリーンアップする必要があります
バックポート	<a href="#">HIVE-22448</a> : CBO: group-by キーを使用して複数のカウントを個別に展開する
バックポート	<a href="#">HIVE-22248</a> : 統計情報が持続する問題を修正
バックポート	<a href="#">HIVE-22210</a> : ベクトル化により、フィルタリングに関係する計算出力列が再利用される可能性がある
バックポート	<a href="#">HIVE-21531</a> : ベクトル化: すべての NULL ハッシュコードが Murmur3 を使用して計算されない
バックポート	<a href="#">HIVE-20419</a> : ベクトル化: ハッシュマップキーで使用された VectorPartitionDesc 後ののミューテーションを防ぐ
バックポート	<a href="#">HIVE-19388</a> : VectorMapJoinCommonOperator 初期化 ClassCastException 中
バックポート	<a href="#">HIVE-21584</a> : Java 11 の準備: システムクラスローダーが URL ではないClassLoader
バックポート	<a href="#">HIVE-25107</a> : クラスパスのロギングは DEBUG レベルにすべき (#2271)
バックポート	<a href="#">HIVE-22097</a> : Java 11 用の互換性のない java.util。ArrayList

型	説明
バックポート	<a href="#">HIVE-23938</a> : LLAP: JDK11 - 一部の GC ログファイルのローテーションに関連する jvm 引数が使用できなくなった
バックポート	<a href="#">HIVE-26226</a> : upgrade-acid で hive メタストアから jdk.tools dep を除外する
バックポート	<a href="#">HIVE-17879</a> : Datanucleus Maven プラグインのアップグレード
バックポート	<a href="#">HIVE-27004</a> : 8 を超える Java バージョンでは DateTimeFormatterBuilder#appendZoneText can not parse 'UTC+'
バックポート	<a href="#">HIVE-16812</a> : VectorizedOrcAcidRowBatchReader 削除イベントをフィルタリングしない
バックポート	<a href="#">HIVE-17917</a> : VectorizedOrcAcidRowBatchReader.computeOffsetAndBucket 最適化
バックポート	<a href="#">HIVE-19985</a> : ACID: 読み取り専用クエリの ROW_ID セクションのデコードをスキップする
バックポート	<a href="#">HIVE-20635</a> : VectorizedOrcAcidRowBatchReader は元のファイルの削除イベントをフィルタリングしません
Upgrade	Javadoc を 3.3.1 にアップグレード
Upgrade	Javassist を 3.24.1-GA にアップグレード
Upgrade	2.0.0-M14 apache-directory-server への更新



## 新しい設定

名前	分類	説明
hive.metastore.fs.drop.partition.threads	hive-site	ドロップパーティションスレッドプール内のコアスレッドの数。
hive.metastore.fs.drop.partition.keepalive.time	hive-site	(スレッドプールからの) アイドル状態のドロップパーティション非同期スレッドが新しいタスクが到着するのを待ってから終了する時間 (秒単位)。
hive.metastore.fs.drop.partition.threadpool.max.queue.size	hive-site	ファイルシステムからパーティションを削除する際にスレッドプール内で使用される最大キューサイズ。
hive.groupby.enable.deterministic.distribution	hive-site	リデューサーへのキーの確定的配布を有効にします。ランダムパーティショニングに使用される rand 関数を呼び出す際に、定数のシード値を渡します。
hive.privilege.synchronizer	hive-site	HiveServer2 で外部オーソライザーからの権限を定期的に同期するかどうか。
hive.cli.query.file.encoding	hive-site	cli 引数で提供されるすべてのタイプのクエリファイル (クエリファイル、init クエリファイル、rc ファイルなど) のファイルエンコーディング。

名前	分類	説明
hive.emr.tez.shuffle.enabled	hive-site	Hive on Tez ジョブは、デフォルトで mapreduce_shuffle の代わりに tez_shuffle をデフォルトのシャッフルハンドラーとして使用するようになりました。

## 非推奨の構成

以下の設定プロパティは [HIVE-23354](#) の結果として廃止され、Amazon EMR リリース 6.11.0 以降ではサポートされなくなりました。

名前	デフォルト値
hive.mapred.reduce.tasks speculative.execution	false
tez.am.speculation.enabled	false

## Amazon EMR 6.10.0 - Hive リリースノート

### Amazon EMR 6.10.0 - Hive の変更点

型	説明
機能	IAM パススルー (HiveCLI /Steps API) を介して Apache Hive クエリ (書き込み) の AWS Lake Formation ベースアクセスコントロールを有効にします。 <a href="#">HiveCLI</a>
改良点	ログサイズを減らすため、デフォルトではコンフィグ hive.log.explain.output を無効にします

型	説明
バックポート	<a href="#">HIVE-26408</a> : ベクトル化: スクラッチ列の割り当て解除を修正し、子を出力 ConstantVectorExpression として再利用しない
バックポート	<a href="#">HIVE-22269</a> : <a href="#">HIVE-20703</a> による統計情報の欠落による動的パーティションの挿入クエリのリデューサーカウントの誤りを修正。
バックポート	<a href="#">HIVE-22891</a> : 非 LLAP CombineHiveRecord 実行モードでの PartitionDesc 抽出をスキップする
バックポート	<a href="#">HIVE-23804</a> : Hive メタストアスキーマに列統計固有のテーブル用のデフォルトデータベースを追加して下位互換性を確保
バックポート	<a href="#">HIVE-25277</a> : コストの高い クラウドオブジェクトストアの Hive パーティションの削除が遅い ListFiles
バックポート	<a href="#">HIVE-19202</a> : isBucketedInput 「」 () NullPointerException のにより CBO が失敗しました HiveAggregate。
バックポート	<a href="#">HIVE-19048</a> : beeline Initscript エラーが無視される問題を修正
バックポート	<a href="#">HIVE-21085</a> : マテリアライズドビューレジストリが外部以外の tez セッションを開始する
バックポート	<a href="#">HIVE-21675</a> : CREATE VIEW IF NOT EXISTS でビューが既に存在する場合「OK」ではなくエラーが返される。これは Hive 2 からのリグレッションです。
バックポート	<a href="#">HIVE-21646</a> : Tez: スレッドログコンテキストからのエスケープ TezTasks を防止

型	説明
バックポート	<a href="#">HIVE-22054</a> : ディレクトリが空かどうかをチェックする再帰的な一覧表示を避ける
バックポート	<a href="#">HIVE-16587</a> : nul値がネストされた複合型を挿入したときの NPE
バックポート	<a href="#">HIVE-22647</a> : デフォルトでセッションプールを有効にする
バックポート	<a href="#">HIVE-13288</a> : DagUtils.localizeResource の例外メッセージの混乱
バックポート	<a href="#">HIVE-23870</a> : WritableHiveCharObjectInspector.getPrimitiveJavaObject / で複数のテキスト変換を最適化する HiveCharWritable
バックポート	<a href="#">HIVE-21498</a> : Thrift を 0.13.0 にアップグレード
バックポート	<a href="#">HIVE-24378</a> : 先頭と末尾のスペースが小数点変換の前に削除されない
バックポート	<a href="#">HIVE-21341</a> : 妥当なデフォルト: hive.server2.idle.operation.timeout と hive.server2.idle.session.timeout が高すぎる
バックポート	<a href="#">HIVE-22465</a> : に ssl conf を追加する TezConfigurationFactory
バックポート	<a href="#">HIVE-24710</a> : count(*) の PTF イテレーションを最適化して CPU と IO のコストを削減
バックポート	<a href="#">HIVE-15406</a> : 新しい「trunc」関数のベクトル化を検討する
バックポート	<a href="#">HIVE-21541</a> : HIVE-15406 から asf ヘッダーが見つからなかった問題を修正
バックポート	<a href="#">HIVE-24808</a> : 解析された日付をキャッシュする

型	説明
バックポート	<a href="#">HIVE-24746</a> : PTF: TimestampValueBoundaryScanner 範囲計算中に最適化可能
バックポート	<a href="#">HIVE-25059</a> : レプリケーション中に変更イベントが名前変更に変換される
バックポート	<a href="#">HIVE-25142</a> : マップ結合高速ハッシュテーブルでの再ハッシュにより、大きなキーが破損する
バックポート	<a href="#">HIVE-23756</a> : package.jdo ファイルにさらに制約を追加しました
バックポート	<a href="#">HIVE-25150</a> : HIVE-24378 の一部として修正されたスペース文字と同様に、10 進数変換の前のタブ文字が削除されない
バックポート	<a href="#">HIVE-25093</a> : date_format() UDF は UTC タイムゾーンのみで出力を返している
バックポート	<a href="#">HIVE-25268</a> : ローカルタイムゾーンが UTC 以外の場合、date_format udf は 1900 年より前の日付に対して間違った結果を返す
バックポート	<a href="#">HIVE-25338</a> : 入力が空の場合の変換 UDF 内の AIOBE
バックポート	<a href="#">HIVE-22400</a> : 時刻を含む UDF 分が NULL を返す
バックポート	<a href="#">HIVE-25058</a> : PTF: TimestampValueBoundaryScanner 範囲計算 pt2 - 中に最適化可能 isDistanceGreater
バックポート	<a href="#">HIVE-25449</a> : datediff() を UTC 以外の一部のタイムゾーンで tez タスクで実行すると、間違った出力が返される

型	説明
バックポート	<a href="#">HIVE-23688</a> : ベクトル化: null IndexArrayOutOfBoundsException 値を含むマップタイプの場合
バックポート	<a href="#">HIVE-22247</a> : パーティションのタスク出力が空 FileNotFoundException になると HiveHFile OutputFormat がスローする
バックポート	<a href="#">HIVE-25570</a> : Hive は、コマンド挿入の上書き場所の承認用に完全な URL パスを送信する必要がある
バックポート	<a href="#">HIVE-22903</a> : パーティション節に定数式がある場合、ベクトル化された row_number() は 1 バッチ後に行番号をリセットする
バックポート	<a href="#">HIVE-25549</a> : PARTITION BY 句または ORDER BY 句に式があるウィンドウ関数の結果が正しくない
バックポート	<a href="#">HIVE-25579</a> : LOAD 上書きは上書きではなく追記される
バックポート	<a href="#">HIVE-25659</a> : IN/(NOT IN) を含むメタストアダイレクト SQL クエリは、SQL DB で許可されている最大パラメータに基づいて分割する必要がある
バックポート	<a href="#">HIVE-20502</a> : カラム統計が使用されている場合に skewjoin_mapjoin10.q を実行しているときの NPE を修正。
バックポート	<a href="#">HIVE-25765</a> : skip.header.line.count プロパティは、ファイルサイズが大きい FetchOperator 場合、の各ブロックの行をスキップします

型	説明
バグ	hive.stats.column.autogather と hive.groupby.skewindata が両方とも有効になっている特定のシナリオでの挿入時の NPE を修正
バグ	mapred.tasktracker.expiry.interval 値が設定されていない場合の NPE を修正

## Amazon EMR 6.9.0 - Hive リリースノート

### Amazon EMR 6.9.0 - Hive の変更点

型	説明
Upgrade	Jetty を <a href="#">9.4.48.v20220622</a> にアップグレード
Upgrade	Hadoop 3.3.3 をサポート
機能	Amazon EMR Hive を Lake Formation と統合して、GCSC API を使用するインタラクティブワークロードを実現します。
機能	Amazon EMR Hive と Iceberg の統合。
改良点	Amazon EMR セキュリティ設定を使用して <a href="#">転送時の暗号化が有効になっている場合は</a> 、HiveServer2 で SSL を有効にします。
改良点	Hive EMRFS Amazon S3 最適化コミッターをデフォルトで有効にします。詳細については、「 <a href="#">Hive EMRFS S3 向けに最適化されたコミッターを有効にする</a> 」を参照してください。
改良点	マップされたバージョンののみを継承HiveHBaseTableInputFormatV2 す

型	説明
	<p>るを追加して InputFormat、<a href="#">SPARK-34210</a>を修正します。これを使うには <code>hive.hbase.inputformat.v2</code> を <code>true</code> に設定してください。</p>
改良点	<p>TezAM を終了してすぐに新しいバージョンを起動するのではなく、<a href="#">hive.cli.tez.session.async</a> を使用してバックグラウンドで起動するのを待ちます。このタイムアウトを秒単位で設定するには <code>hive.emr.cli.tez.session.open.timeout</code> を使用します。</p>
改良点	<p><a href="#">hive.conf.restricted.list.append</a> オプションを追加すると、カンマで区切られた設定を既存の制限付き設定リスト <code>hive.conf.restricted.list</code> に追加できます。</p>
改良点	<p>データベースのロケーションが定義されていないために Hive クエリが失敗したときのエラーメッセージがわかりやすくなりました。</p>
バックポート	<p><a href="#">HIVE-24484</a>: Hadoop を 3.3.1 に、Tez を 0.10.2 にアップグレード</p>
バックポート	<p><a href="#">HIVE-22398</a>: 経由で YARN キュー管理を削除します ShimLoader。</p>
バックポート	<p><a href="#">HIVE-23190</a>: LLAP: ファイルシステムオブジェクト <code>IndexCache</code> を <code>TezSpillRecord</code> に渡すように変更します</p>
バックポート	<p><a href="#">HIVE-22185</a>: HADOOP-15832 は、MiniYarn クラスタを使用したテストで問題を引き起こします。</p>
バックポート	<p><a href="#">HIVE-21670</a>: <code>mockito-all</code> を <code>mockito-core</code> 依存関係に置き換え。</p>



型	説明
バックポート	<a href="#">HIVE-24542</a> : アップグレード用に Guava を準備。
バックポート	<a href="#">HIVE-23751</a> : QTest: HADOOP-16582 の後に ProxyFileSystem で #mkdirs() メソッドをオーバーライドして位置合わせする。
バックポート	<a href="#">HIVE-21603</a> : Java 11 の準備: powermock のバージョンを更新。
バックポート	<a href="#">HIVE-24083</a> : Hadoop 3.3.0 の hcatalog エラー: 認証タイプが必要。
バックポート	<a href="#">HIVE-24282</a> : 明示的に指定されない限り、列の表示で出力列をソートしないでください。
バックポート	<a href="#">HIVE-20656</a> : 妥当なデフォルト: マップ集約メモリの設定が重度すぎる。
バックポート	<a href="#">HIVE-25443</a> : 1024 を超える値がある場合、Arrow は複雑なデータ型をシリアル化/逆シリアル化 SerDe できない
バックポート	<a href="#">HIVE-19792</a> : orc を 1.5.2 にアップグレードして、decimal_64 スキーマ進化テストを有効化する。
バックポート	<a href="#">HIVE-20437</a> : 浮動小数点数、倍精度および 10 進数からのスキーマの進化を処理する。
バックポート	<a href="#">HIVE-21987</a> : Hive は、10 進数のアノテーションが付けられた Parquet int32 を読み取ることができない。
バックポート	<a href="#">HIVE-20038</a> : バケット化されていないテーブルやパーティション化されたテーブルの更新クエリで NPE がスローされる。

## Amazon EMR 6.9.0 - Hive の既知の問題

- Amazon EMR 6.6.0 から 6.9.x では、動的パーティションと ORDER BY 句または SORT BY 句を使用した INSERT クエリには常に 2 つのリデューサーがあります。この問題は、OSS が変更された [HIVE-20703](#) が原因です。これにより、動的ソートパーティションの最適化がコストベースの決定下に置かれます。ワークロードで動的パーティションのソートが不要な場合は、hive.optimize.sort.dynamic.partition.threshold プロパティを -1 に設定して新機能を無効にし、リデューサーの数を正しく計算することをお勧めします。この問題は、[HIVE-22269](#) の一部として OSS Hive で修正され、Amazon EMR 6.10.0 で修正されています。

## Amazon EMR 6.8.0 - Hive リリースノート

### Amazon EMR 6.8.0 - Hive の変更点

型	説明
改良点	msck コマンドのファイルシステム呼び出しを減らします。パフォーマンスの向上 (10,000 以上のパーティションで最大 15~20 倍)
バックポート	<a href="#">HIVE-20678</a> : HiveHBaseTableOutputFormat は互換性 HiveOutputFormat を確保するためにを実装する必要があります
バックポート	<a href="#">HIVE-21040</a> : msck はディレクトリツリーの最後のレベルで不要なファイルを一覧表示する
バックポート	<a href="#">HIVE-21460</a> : データをロードした後に select * クエリを実行すると、誤った結果になる
バックポート	<a href="#">HIVE-21660</a> : union all の後で view with explode を使用すると間違った結果になる
バックポート	<a href="#">HIVE-22505</a> : ベクトル化された演算子の選択が間違っ て ClassCastException いる

型	説明
バックポート	<a href="#">HIVE-22513</a> : フィルター操作でのキャスト列の定数伝播により、誤った結果が生じる可能性がある
バックポート	<a href="#">HIVE-23435</a> : 完全外部結合の結果で行が欠落している
バックポート	<a href="#">HIVE-24209</a> : ベクトル化が有効な場合の NOT BETWEEN 操作の検索指数の変換が正しくない
バックポート	<a href="#">HIVE-24934</a> : GenericUDFSQCountCheck では VectorizedExpressions 注釈は必要ありません
バックポート	<a href="#">HIVE-25278</a> : ウィンドウ式で無効な変換を行う HiveProjectJoinTransposeRule 可能性がある
バックポート	<a href="#">HIVE-25505</a> : 最初の行が空白の場合、ヘッダーの skip.header.line.count で不正な結果になる
バックポート	<a href="#">HIVE-26080</a> : accumulo-core を 1.10.1 にアップグレード
バックポート	<a href="#">HIVE-26235</a> : バイナリ列の OR 条件が空の結果を返す
バグ	起動時の stderr の複数の SLF4J バインディング警告ログを修正
バグ	パーティションとテーブルが異なるファイルシステムにある場合に、SHOW TABLE EXTENDED クエリが Wrong FS エラーで失敗する問題を修正。

## Amazon EMR 6.8.0 - Hive の既知の問題

- Amazon EMR 6.6.0 から 6.9.x では、動的パーティションと ORDER BY 句または SORT BY 句を使用した INSERT クエリには常に 2 つのリデューサーがあります。この問題は、OSS が変更された [HIVE-20703](#) が原因です。これにより、動的ソートパーティションの最適化がコストベースの決定下に置かれます。ワークロードで動的パーティションのソートが不要な場合は、hive.optimize.sort.dynamic.partition.threshold プロパティを -1 に設定して新機能を無効にし、リデューサーの数を正しく計算することをお勧めします。この問題は、[HIVE-22269](#) の一部として OSS Hive で修正され、Amazon EMR 6.10.0 で修正されています。

## Amazon EMR 6.7.0 - Hive リリースノート

### Amazon EMR 6.7.0 - Hive の変更点

型	説明
機能	<a href="#">Amazon EMR Hive との統合 LakeFormation。</a>
機能	Hive EMRFS Amazon S3 最適化コミッター用の監査ログが追加されました。Hive 設定: <code>hive.blobstore.output-committer.logging</code> 、デフォルト: <code>false</code>
機能	挿入時に削除されたターゲットディレクトリは、空の選択結果でパーティション化されていないテーブル/静的パーティションに上書きされ、Hive 2.x と同様に動作するようになりました。Hive 設定: <code>hive.emr.iow.clean.target.dir</code> 、デフォルト: <code>false</code>
バグ	Hive EMRFS Amazon S3 最適化コミッターをパーティションバケットソートで使用すると断続的に発生するクエリエラーを修正しました。
Upgrade	Hive をバージョン 3.1.3 にアップグレードしました 詳細については、「 <a href="#">Apache Hive 3.1.3 リリースノート</a> 」を参照してください。

型	説明
Upgrade	Parquet を <a href="#">1.12.2</a> にアップグレードしました。
バックポート	<a href="#">HIVE-20065</a> : メタストアを jackson 1.x に依存することはできない
バックポート	<a href="#">HIVE-20071</a> : jackson 2.x に移行して使用を防止する
バックポート	<a href="#">HIVE-20607</a> : TxnHandler PreparedStatement を使用して直接 SQL クエリを実行する必要があります
バックポート	<a href="#">HIVE-20740</a> : ObjectStore.setConf メソッドでグローバルロックを削除する
バックポート	<a href="#">HIVE-20961</a> : NVL の実装を廃止
バックポート	<a href="#">HIVE-22059</a> : hive-exec jar には (fasterxml) jackson ライブラリが含まれていない
バックポート	<a href="#">HIVE-22351</a> : でのスレッド ObjectStoreの誤った使用を修正 TestObjectStore
バックポート	<a href="#">HIVE-23534</a> : メッセージなしで をキャッチ MetaException するときの RetryingMetaStoreClient#invoke の NPE
バックポート	<a href="#">HIVE-24048</a> : Jackson のコンポーネントをバージョン 2.10 の最新版 (Hive) に一致させる
バックポート	<a href="#">HIVE-24768</a> : バージョン交換にはどこでも jackson-bom を使用する
バックポート	<a href="#">HIVE-24816</a> : CVE-2020-25649 により、jackson を 2.10.5.1 または 2.11.0+ にアップグレードする

型	説明
バックポート	<a href="#">HIVE-25971</a> : キャッシュされたスレッドプールが閉じられていないため、Tez タスクのシャットダウンが遅れる
バックポート	<a href="#">HIVE-26036</a> : の <code>getMTable ()</code> によって引き起こされる NPE ObjectStore

## Amazon EMR 6.7.0 - Hive の既知の問題

- `join` と同じ列にウィンドウ関数があるクエリは、[HIVE-25278](#) で報告されているように無効な変換を引き起こし、誤った結果やクエリ失敗の原因となることがあります。回避策は、このようなクエリのクエリレベルで CBO を無効にすることです。この修正は、6.7.0 以降の Amazon EMR リリースで利用できるようになります。詳細については、AWS サポートにお問い合わせください。
- Amazon EMR 6.6.0 から 6.9.x では、動的パーティションと `ORDER BY` 句または `SORT BY` 句を使用した `INSERT` クエリには常に 2 つのリデューサーがあります。この問題は、OSS が変更された [HIVE-20703](#) が原因です。これにより、動的ソートパーティションの最適化がコストベースの決定下に置かれます。ワークロードで動的パーティションのソートが不要な場合は、`hive.optimize.sort.dynamic.partition.threshold` プロパティを `-1` に設定して新機能を無効にし、リデューサーの数を正しく計算することをお勧めします。この問題は、[HIVE-22269](#) の一部として OSS Hive で修正され、Amazon EMR 6.10.0 で修正されています。

## Amazon EMR 6.6.0 - Hive リリースノート

### Amazon EMR 6.6.0 - Hive の変更点

型	説明
Upgrade	Parquet を <a href="#">1.12.1</a> にアップグレード。
Upgrade	jetty jars バージョンを 9.4.43.v20210629 にアップグレード

型	説明
バグ	Hive クラスターで LLAP が有効になっていると、すべてのタスク/コアノードに Hive がインストールされる問題を修正しました。
バックポート	<a href="#">HIVE-25942</a> : CVE-2021-29425 により、commons-io を 2.8.0 にアップグレード
バックポート	<a href="#">HIVE-25726</a> : CVE-2020-13936 により、velocity を 2.3 にアップグレード
バックポート	<a href="#">HIVE-25680</a> : #get_table_meta HiveMetastore Server API に任意の HiveMetastore 認証モデルの使用を許可します。
バックポート	<a href="#">HIVE-25554</a> : arrow バージョンを 0.15 にアップグレード
バックポート	<a href="#">HIVE-25242</a> : vectorized.adaptor = を選択すると、クエリの実行が非常に遅くなる
バックポート	<a href="#">HIVE-25085</a> : MetaStore クライアントはセッション間で共有されなくなりました。
バックポート	<a href="#">HIVE-24827</a> : Hive 集約クエリは、テキスト以外のファイルに対して誤った結果を返します。
バックポート	<a href="#">HIVE-24683</a> : Hadoop23Shims getFileId は、存在しないパスに対して NPE が発生しやすい
バックポート	<a href="#">HIVE-24656</a> : マップ型および配列型で is null を使用したクエリの CBO が失敗する
バックポート	<a href="#">HIVE-24556</a> : 孫がないケース DefaultGraphWalker 向けに最適化
バックポート	<a href="#">HIVE-24408</a> : Parquet を 1.11.1 にアップグレード

型	説明
バックポート	<a href="#">HIVE-24391</a> : branch-3.1 の FIX TestOrcFile 障害を修正
バックポート	<a href="#">HIVE-24362</a> : ノード数が多いツリーでは AST ツリー処理が最適ではない
バックポート	<a href="#">HIVE-24316</a> : branch-3.1 で ORC を 1.5.6 から 1.5.8 にアップグレード
バックポート	<a href="#">HIVE-24307</a> : property-file と -e パラメータで Beeline が失敗する
バックポート	<a href="#">HIVE-24245</a> : カウントと個別のオーバーパーティションを含むベクトル化された PTF が誤った結果を生成する。
バックポート	<a href="#">HIVE-24224</a> : 圧縮ファイルで Hive on Tez のヘッダー/フッターがスキップされる問題を修正
バックポート	<a href="#">HIVE-24157</a> : CAST タイムスタンプ ↔ 数値でストリクトモードが失敗する
バックポート	<a href="#">HIVE-24113</a> : GenericUDFToUnixTimeStamp の NPE
バックポート	<a href="#">HIVE-23987</a> : arrow をバージョン 0.11.0 にアップグレード
バックポート	<a href="#">HIVE-23972</a> : 外部クライアント ID を LLAP 外部クライアントに追加
バックポート	<a href="#">HIVE-23806</a> : スキーマが拡張された場合に備えて、すべてのパーティションの列統計状態の消去を回避する。これにより、alter table add columns ステートメントの実行時間が改善される。



型	説明
バックポート	<a href="#">HIVE-23779</a> : BasicStatsTask Info が beeline コンソールに出力されない
バックポート	<a href="#">HIVE-23306</a> : System.getProperty によって設定されている構成がある場合、RESET コマンドが機能しない
バックポート	<a href="#">HIVE-23164</a> : デーモンスレッド以外のスレッドが原因でサーバーが正しく終了しない
バックポート	<a href="#">HIVE-22967</a> : Hive on Tez の hive.reloadable.aux.jars.path のサポート
バックポート	<a href="#">HIVE-22934</a> : エラーストリームへの Hive サーバーのインタラクティブログカウンター
バックポート	<a href="#">HIVE-22901</a> : 変数置換により循環参照で OOM が発生する可能性がある
バックポート	<a href="#">HIVE-22769</a> : 圧縮テキストファイルの分割生成中にクエリ結果が不正確になり、クエリが失敗する
バックポート	<a href="#">HIVE-22716</a> : への読み取り ByteBuffer が壊れている ParquetFooterInputFromCache
バックポート	<a href="#">HIVE-22648</a> : Parquet を 1.11.0 にアップグレード
バックポート	<a href="#">HIVE-22640</a> : Decimal64ColumnVector: パーティション列タイプが 10 進数 ClassCastException の場合
バックポート	<a href="#">HIVE-22621</a> : 不安定なテストケース: TestLlapSignerImpl.testSigning

型	説明
バックポート	<a href="#">HIVE-22533</a> : LLAP デーモンのウェブ UI の潜在的な脆弱性を修正
バックポート	<a href="#">HIVE-22532</a> : PTF" がランク/DenseRank 機能を通じて制限を誤ってプッシュすることがある
バックポート	<a href="#">HIVE-22514</a> : 大量のメモリを消費 HiveProto LoggingHook する可能性がある
バックポート	<a href="#">HIVE-22476</a> : hive.fetch.task.converion が none に設定されていると、Hive の datediff 関数が一貫性のない結果を返した
バックポート	<a href="#">HIVE-22429</a> : hive 3 で bucketing_version 1 を使用して移行されたクラスター化されたテーブルは、挿入に bucketing_version 2 を使用する
バックポート	<a href="#">HIVE-22412</a> : 説明時に NPE を StatsUtils スローする
バックポート	<a href="#">HIVE-22360</a> : ロードされたファイルにテーブルスキーマ内の列よりも多くの列がある場合、最後の列に間違った結果 MultiDelimitSerDe を返します
バックポート	<a href="#">HIVE-22332</a> : ORC-540 以降、Hive は有効なスキーマ展開設定を確認する必要がある
バックポート	<a href="#">HIVE-22331</a> : 引数なしの unix_timestamp は、タイムスタンプを秒単位ではなくミリ秒単位で返す
バックポート	<a href="#">HIVE-22275</a> : OperationManager.queryIdOperation が複数の queryIdsを適切にクリーンアップしない

型	説明
バックポート	<a href="#">HIVE-22273</a> : 一時ディレクトリを削除するとアクセスチェックが失敗する
バックポート	<a href="#">HIVE-22270</a> : commons-io を 2.6 にアップグレード
バックポート	<a href="#">HIVE-22241</a> : UDF を実装して、内部表現とグレゴリオ暦とユリウス暦のハイブリッド暦を使用して日付/タイムスタンプを解釈する
バックポート	<a href="#">HIVE-22241</a> : UDF を実装して、内部表現とグレゴリオ暦とユリウス暦のハイブリッドを使用し日付/タイムスタンプを解釈する
バックポート	<a href="#">HIVE-22232</a> : hive.order.columnalignment が false に設定されている場合の NPE
バックポート	<a href="#">HIVE-22231</a> : Knox 経由の大きなサイズの Hive クエリが、パイプが壊れて書き込みの失敗で失敗する
バックポート	<a href="#">HIVE-22221</a> : 外部クライアントをラップする - LlapBaseInputFormat#getSplits を減らす必要がある
バックポート	<a href="#">HIVE-22208</a> : マスク列のあるテーブルの結合を含むクエリを書き換えると、予約キーワードの付いた列名がエスケープ解除される
バックポート	<a href="#">HIVE-22197</a> : Common Merge join でクラスキャスト例外が発生する。
バックポート	<a href="#">HIVE-22170</a> : from_unixtime と unix_timestamp はユーザーセッションのタイムゾーンを使用する必要がある

型	説明
バックポート	<a href="#">HIVE-22169</a> : Tez: Tez に存在しない計画ファイルを探 SplitGenerator そうとします
バックポート	<a href="#">HIVE-22168</a> : LLAP キャッシュのホットパスから非常にコストのかかるロギングを削除
バックポート	<a href="#">HIVE-22161</a> : org.apache.hadoop.hive ql.udf.UDFType クラスで UDF: FunctionRegistry 同期
バックポート	<a href="#">HIVE-22120</a> : 特定の境界条件での左外部マップ結合の誤った結果/ArrayOutOfBound例外を修正
バックポート	<a href="#">HIVE-22115</a> : プロパティが false に設定されている場合、クエリルーティングアペンダーが作成されないようにする
バックポート	<a href="#">HIVE-22113</a> : AMReporter 関連の LLAP シャットダウンの防止 RuntimeException
バックポート	<a href="#">HIVE-22106</a> : パーティション評価用のクロスクエリ同期を削除
バックポート	<a href="#">HIVE-22099</a> : HIVE-20007 以降、いくつかの日付関連の UDF がユリウス暦を適切に処理できない
バックポート	<a href="#">HIVE-22037</a> : OOM が原因でシャットダウンするときに HS2 がログを記録する必要がある
バックポート	<a href="#">HIVE-21976</a> : Calcite のオフセットはゼロではなく null にする必要があります HiveSortLimit
バックポート	<a href="#">HIVE-21924</a> : ヘッダー/フッターが存在していてもテキストファイルを分割する

型	説明
バックポート	<a href="#">HIVE-21913</a> : GenericUDTFGetSplits は LLAP と同じ方法でユーザー名を処理する必要があります
バックポート	<a href="#">HIVE-21905</a> : FetchOperator クラス全体のジェネリックの改善
バックポート	<a href="#">HIVE-21902</a> : HiveServer2 UI: Jetty レスポンスヘッダーには X-Frame-Options が必要です
バックポート	<a href="#">HIVE-21888</a> : hive.parquet.timestamp.skip.conversion のデフォルトを true に設定
バックポート	<a href="#">HIVE-21868</a> : CAST...FORMAT をベクトル化
バックポート	<a href="#">HIVE-21864</a> : LlapBaseInputFormat#closeAll
バックポート	<a href="#">HIVE-21863</a> : WHEN 式の Vectorizer 型キャストを改善
バックポート	<a href="#">HIVE-21862</a> : ORC ppd がタイムスタンプ付きの間違った結果を生成する
バックポート	<a href="#">HIVE-21846</a> : 定期的に LlapDaemon メトリクスを取得するスレッドを TezAM に作成する
バックポート	<a href="#">HIVE-21837</a> : 選択した列に完全に null 値がある場合、例外を MapJoin スローしています
バックポート	<a href="#">HIVE-21834</a> : 不要な呼び出しを避けてフィルター条件を簡素化
バックポート	<a href="#">HIVE-21832</a> : キュー/サービング/レスポンスの平均時間を取得する新しいメトリクス
バックポート	<a href="#">HIVE-21827</a> : の SemanticAnalyzer 複数の呼び出しが getTableObjectByName メソッドを経由しない

型	説明
バックポート	<a href="#">HIVE-21822</a> : 新しい API メソッドを使用して LlapDaemon メトリクスを公開する
バックポート	<a href="#">HIVE-21818</a> : CBO: コピーに TableRelOptHiveTableメタストアトラフィックがある
バックポート	<a href="#">HIVE-21815</a> : ORC ファイル内の統計が 2 回解析される
バックポート	<a href="#">HIVE-21805</a> : HiveServer2: 高速 ShutdownHookManager APIs
バックポート	<a href="#">HIVE-21799</a> : 結合キーが集計列にある場合 DynamicPartitionPruningOptimization、NullPointerException で
バックポート	<a href="#">HIVE-21794</a> : マテリアライズドビューパラメータを に追加する sqlStdAuthSafeVarNameRegexes
バックポート	<a href="#">HIVE-21768</a> : JDBC: 囲まれていない UNION クエリのデフォルトのユニオンプレフィックスを削除
バックポート	<a href="#">HIVE-21746</a> : ArrayIndexOutOfBoundsException 動的にパーティション分割されたハッシュ結合中、CBO が無効
バックポート	<a href="#">HIVE-21717</a> : 移動タスクでディレクトリの名前変更失敗する。
バックポート	<a href="#">HIVE-21685</a> : 複数の IN 句を含むクエリの簡略化が間違っている
バックポート	<a href="#">HIVE-21681</a> : 書式設定された説明で複数のプライマリキーについて、誤った情報が表示される

型	説明
バックポート	<a href="#">HIVE-21651</a> : protobuf serde を hive-exec に移動。
バックポート	<a href="#">HIVE-21619</a> : SQL explain 拡張でタイムスタンプ型を精度なしで出力する
バックポート	式に CONCAT が含まれている場合、 <a href="#">HIVE-21592</a> : OptimizedSql は表示されません
バックポート	<a href="#">HIVE-21576</a> : CAST...FORMAT と SQL:2016 日時フォーマットの限定リストを導入
バックポート	<a href="#">HIVE-21573</a> : 認証が delegationToken に設定されている場合、バイナリトランスポートはプリンシパルを無視する
バックポート	<a href="#">HIVE-21550</a> : TestObjectStore テストがフラッキー - リクエストされた時間内にロックを取得できませんでした
バックポート	<a href="#">HIVE-21544</a> : 定数伝播により折りたたみ中に coalesce/case/when 式が破損する
バックポート	<a href="#">HIVE-21539</a> : GroupBy + 同じ列の 句が原因でクエリの書き換えが正しくない
バックポート	<a href="#">HIVE-21538</a> : Beeline: コンソールリーダー経由でパスワードソースが接続パラメータに渡されなかった
バックポート	<a href="#">HIVE-21509</a> : LLAP が破損した列ベクトルをキャッシュし、間違ったクエリ結果を返すことがある
バックポート	<a href="#">HIVE-21499</a> : コマンドの作成が で失敗した場合、 はレジストリから関数を削除しないでください AlreadyExistsException

型	説明
バックポート	<a href="#">HIVE-21496</a> : 順序付けられていないバッファの自動サイズ設定はオーバーフローする可能性がある
バックポート	<a href="#">HIVE-21468</a> : JDBC ストレージハンドラーの識別子名の大文字と小文字の区別
バックポート	<a href="#">HIVE-21467</a> : 廃止予定の junit.framework.Assert インポートを削除
バックポート	<a href="#">HIVE-21435</a> : 構築中に TASK_ATTEMPT_ID conf からタスク番号を取得 LlapBasel nputFormat する必要がある SubmitWorkRequestProto
バックポート	<a href="#">HIVE-21389</a> : HIVE-21247 以降、Hive のデストリビューションに javax.ws.rs-api.jar が見つからない
バックポート	<a href="#">HIVE-21385</a> : 分割不可能な計算の JDBC ソースへのプッシュダウンを無効にできるようにする
バックポート	<a href="#">HIVE-21383</a> : JDBC ストレージハンドラー: 指定されている場合、カタログとスキーマを使用してテーブルを取得する
バックポート	<a href="#">HIVE-21382</a> : Group by キー削減の最適化 - query23 でキーが削減されない
バックポート	<a href="#">HIVE-21362</a> : protobuf ファイルから読み込むための入力フォーマットと serde を追加。
バックポート	<a href="#">HIVE-21340</a> : CBO: 非キー列を にフィードするプルーニング SemiJoin
バックポート	<a href="#">HIVE-21332</a> : ロックされたバッファの代わりにロックされていないバッファを消去する



型	説明
バックポート	<a href="#">HIVE-21329</a> : オペレータパイプラインに依存するカスタム Tez ランタイムの順序付けされていない出力バッファサイズ
バックポート	<a href="#">HIVE-21295</a> : Hive StorageHandler 規則を使用して日付を文字列に変換する
バックポート	<a href="#">HIVE-21294</a> : ベクトル化: 1-reducer Shuffle はオブジェクトハッシュ関数をスキップできる
バックポート	<a href="#">HIVE-21255</a> : QueryConditionBuilder で を削除する JdbcStorageHandler
バックポート	<a href="#">HIVE-21253</a> : JDBC での DB2 のサポート StorageHandler
バックポート	<a href="#">HIVE-21232</a> : LLAP: キャッシュミスに対応したスプリットアフィニティプロバイダーの追加
バックポート	<a href="#">HIVE-21214</a> : MoveTask : ファイルの重複排除にファイルサイズの代わりに attemptId を使用する compareTempOrDuplicateFiles
バックポート	<a href="#">HIVE-21184</a> : 説明を追加し、コスト情報を含むフォーマット済み CBO プランを説明する
バックポート	<a href="#">HIVE-21182</a> : 計画中に hive スクラッチディレクトリのセットアップをスキップする
バックポート	<a href="#">HIVE-21171</a> : RPC がオンの場合、tez のスクラッチディレクトリの作成をスキップする
バックポート	<a href="#">HIVE-21126</a> : LlapBaseInputFormat#getSplit でセッションレベルのクエリを許可する

型	説明
バックポート	<a href="#">HIVE-21107</a> : 動的にパーティション化されたハッシュ結合中に「Cannot find field」というエラーが発生する
バックポート	<a href="#">HIVE-21061</a> : CTAS クエリが空のソース <code>IllegalStateException</code> に対してで失敗する
バックポート	<a href="#">HIVE-21041</a> : 論理プランからスキーマを取得する <code>ParseException NPE</code>
バックポート	<a href="#">HIVE-21013</a> : Oracle でパーティション列を見つけ <code>JdbcStorageHandler</code> られない
バックポート	<a href="#">HIVE-21006</a> : 再利用の機会がある場合にセミ結合を削除 <code>SharedWorkOptimizer</code> するために拡張する
バックポート	<a href="#">HIVE-20992</a> : <code>hive.metastore.dbaccess.ssl.properties</code> の設定をより意味のある設定に分割する
バックポート	<a href="#">HIVE-20989</a> : JDBC - <code>GetOperationStatus</code> + ログはスリープ経由でクエリの進行をブロックできます
バックポート	<a href="#">HIVE-20988</a> : プライマリキーが複数の列にある <code>group by</code> クエリの結果が正しくない
バックポート	<a href="#">HIVE-20985</a> : <code>select</code> 演算子の入力が一時的な列の場合、ベクトル化によってその一部が出力として再利用されることがある
バックポート	<a href="#">HIVE-20978</a> : " <code>hive.jdbc.*</code> " はに追加する必要があります <code>sqlStdAuthSafeVarNameRegexes</code>
バックポート	<a href="#">HIVE-20953</a> : 作成時にメタストアに追加できない関数を関数レジストリから削除する。

型	説明
バックポート	<a href="#">HIVE-20952</a> : VectorizationContext.java のクリーニング
バックポート	<a href="#">HIVE-20951</a> : LLAP: Xms を常に 50% に設定する
バックポート	<a href="#">HIVE-20949</a> : 物理プランニングにおける PKFK 基数の推定の改善
バックポート	<a href="#">HIVE-20944</a> : クエリのコンパイル中に統計を検証しない
バックポート	<a href="#">HIVE-20940</a> : Calcite のタイプ解像度が Hive よりも厳しいブリッジケース。
バックポート	<a href="#">HIVE-20937</a> : Postgres の JDBC クエリが「LIMIT must not be negative」で失敗する
バックポート	<a href="#">HIVE-20926</a> : ブルームフィルターのエントリが多い場合や統計がない場合、セミ結合削減ヒントが失敗する
バックポート	<a href="#">HIVE-20920</a> : SQL 制約を使用して結合順序変更アルゴリズムを改善
バックポート	<a href="#">HIVE-20918</a> : Calcite から JDBC 接続への計算のプッシュダウンを有効/無効にするフラグ
バックポート	<a href="#">HIVE-20915</a> : 動的ソートパーティションの最適化を HoS と MR で利用できるようにする
バックポート	<a href="#">HIVE-20910</a> : 動的パーティションソートの最適化により、バケットテーブルへの挿入が失敗する

型	説明
バックポート	<a href="#">HIVE-20899</a> : LLAP YARN サービスの Keytab URI は HDFS のみをサポートするように制限されている
バックポート	<a href="#">HIVE-20898</a> : 時間関連の関数では、引数を NULL 値を許容しない型にキャストできない場合がある
バックポート	<a href="#">HIVE-20881</a> : 定数伝播は投影を単純化しすぎる
バックポート	<a href="#">HIVE-20880</a> : hive.stats.filter.in.min.ratio のデフォルト値を更新
バックポート	<a href="#">HIVE-20873</a> : にマーマーハッシュ VectorHashKeyWrapperTwoLong を使用してハッシュの衝突を減らす
バックポート	<a href="#">HIVE-20868</a> : で が子オペレーションを実行している場合、SMB 結合 TezDummyOperator は断続的に失敗 getFinalOp する MapRecord Processor
バックポート	<a href="#">HIVE-20853</a> : llap デーモン API で ShuffleHandler.registerDag を公開する
バックポート	<a href="#">HIVE-20850</a> : 可能であれば条件付きでケースをプロジェクションからディメンションテーブルにプッシュする
バックポート	<a href="#">HIVE-20842</a> : HIVE-20660 で導入されたロジックを group by の統計を推定するように修正
バックポート	<a href="#">HIVE-20839</a> : 動的にパーティション化されたハッシュ結合中に「Cannot find field」というエラーが発生する

型	説明
バックポート	<a href="#">HIVE-20835</a> : 制約と MV の書き換えが相互作用して、Calcite プランナーでループが発生することがある
バックポート	<a href="#">HIVE-20834</a> : キャッシュされたクエリ SemanticAnalyzer からへの参照を保持する Hive QueryResultCache エントリ
バックポート	<a href="#">HIVE-20830</a> : 場合によっては JdbcStorageHandler 範囲クエリアサーションの失敗
バックポート	<a href="#">HIVE-20829</a> : JdbcStorageHandler 範囲分割が NPE をスローする
バックポート	<a href="#">HIVE-20827</a> : 空の配列の結果に一貫性がない
バックポート	<a href="#">HIVE-20826</a> : 左側の結合 + グループを左半結合に変換する HiveSemiJoin ルールを強化
バックポート	<a href="#">HIVE-20821</a> : SUM0 を SUM + COALESCE の組み合わせに書き換える
バックポート	<a href="#">HIVE-20815</a> : JdbcRecordReader.next は例外を食べない
バックポート	<a href="#">HIVE-20813</a> : udf to_epoch_milli はタイムゾーンなしのタイムスタンプもサポートする必要がある。
バックポート	<a href="#">HIVE-20804</a> : 制約付き group by 最適化のさらなる改善
バックポート	<a href="#">HIVE-20792</a> : ゾーン付きのタイムスタンプを挿入するとデータが切り捨てられる

型	説明
バックポート	<a href="#">HIVE-20788</a> : 拡張された SJ の削減では、フィルターを作成すると列が誤ってバックトラックされることがある
バックポート	<a href="#">HIVE-20778</a> : プラン内のすべての結合が非相関ロジックで作成されている場合、結合の順序変更がトリガーされないことがある
バックポート	<a href="#">HIVE-20772</a> : タスクごとの CPU カウンターを LLAP に記録する
バックポート	<a href="#">HIVE-20768</a> : タンプリングウィンドウ UDF の追加
バックポート	<a href="#">HIVE-20767</a> : 結合演算子間の複数のプロジェクトが、制約を使用した結合の順序変更に影響することがある
バックポート	<a href="#">HIVE-20762</a> : NOTIFICATION_LOG のクリーンアップ間隔が 60 秒でハードコードされており、短すぎる
バックポート	<a href="#">HIVE-20761</a> : notification_sequence テーブルの更新対象として選択すると再試行間隔があり、再試行回数が少なすぎる
バックポート	<a href="#">HIVE-20751</a> : arrow をバージョン 0.10.0 にアップグレード
バックポート	<a href="#">HIVE-20746</a> : HiveProtoHookLogger 1 日の終わりにファイルを閉じません。
バックポート	<a href="#">HIVE-20744</a> : SQL 制約を使用して結合順序変更アルゴリズムを改善

型	説明
バックポート	<a href="#">HIVE-20740</a> : ObjectStore.setConf メソッドでグローバルロックを削除します。このチェリーピックは、Hive 3.2 および 4.x 用の HIVE-20740 を 3.1.x にバックポートします。
バックポート	<a href="#">HIVE-20734</a> : Beeline: beeline-site.xml で hive CLI が beeline にリダイレクトする場合、プロンプトの代わりにシステムユーザー名/ダミーパスワードを使用する必要がある
バックポート	<a href="#">HIVE-20731</a> : の JdbcStorageHandler キーストアファイルは承認される必要があります
バックポート	<a href="#">HIVE-20720</a> : JDBC ハンドラーにパーティション列オプションを追加
バックポート	<a href="#">HIVE-20719</a> : hive.optimize.sort.dynamic.partition の最適化とベクトル化をオンにすると、UPDATE の後の SELECT ステートメントが失敗する
バックポート	<a href="#">HIVE-20718</a> : 制約付きの perf CLI ドライバーを追加
バックポート	<a href="#">HIVE-20716</a> : hive.cbo.stats.correlated.multi.key.joins のデフォルト値を true に設定
バックポート	<a href="#">HIVE-20712</a> : HivePointLookupOptimizer デイープケースを抽出する必要がある
バックポート	<a href="#">HIVE-20710</a> : 定数を折りたたみでは、型のない NULL 定数が作成できない場合がある
バックポート	<a href="#">HIVE-20706</a> : external_jdbc_table2.q が断続的に失敗する

型	説明
バックポート	<a href="#">HIVE-20704</a> : 他の関数をサポートする HivePreFilteringRule ように拡張
バックポート	<a href="#">HIVE-20703</a> : 動的ソートパーティションの最適化をコストベースの決定下に置く
バックポート	<a href="#">HIVE-20702</a> : マップ結合選択時のデータ構造を考慮した推定によるオーバーヘッドを考慮する
バックポート	<a href="#">HIVE-20692</a> : NOT x IS (NOT) [TRUE FALSE] 式の折りたたみを有効にする
バックポート	<a href="#">HIVE-20691</a> : org.apache.hadoop.hive.cli.TestMiniLlapCliDriver.testCliDriver[cttl] を修正
バックポート	<a href="#">HIVE-20682</a> : 共有 sessionHive がマスタースレッドによって閉じられると、非同期クエリの実行が失敗する可能性がある
バックポート	<a href="#">HIVE-20676</a> : HiveServer2: PrivilegeSynchronizer がデーモンステータスに設定されていない
バックポート	<a href="#">HIVE-20660</a> : 行の総数をソーステーブルに限定することで、Group by 統計の推定が改善される可能性がある
バックポート	<a href="#">HIVE-20652</a> : 2 つの異なるデータソースの jdbc ドライバーへの JdbcStorageHandler プッシュ結合
バックポート	<a href="#">HIVE-20651</a> : JdbcStorageHandler password は暗号化する必要があります
バックポート	<a href="#">HIVE-20649</a> : Orc 書き込み用の LLAP 対応メモリマネージャー



型	説明
バックポート	<a href="#">HIVE-20648</a> : LLAP: Vector group by 演算子は エグゼキューターごとにメモリを使用する必要がある
バックポート	<a href="#">HIVE-20646</a> : パーティションフィルター条件に IS NOT NULL が設定されている場合、メタス トクエリにプッシュダウンされない
バックポート	<a href="#">HIVE-20644</a> : Hive ランタイム例外による機密 情報の公開を回避する
バックポート	<a href="#">HIVE-20636</a> : outer join 後のヌル値数の推定を 改善
バックポート	<a href="#">HIVE-20632</a> : クエリされたテーブルにマテリ アライズドビューが作成されると、get_splits UDF によるクエリが失敗する
バックポート	<a href="#">HIVE-20627</a> : 同時非同期クエリが断続的に失敗 LockException し、メモリリークが発生する
バックポート	<a href="#">HIVE-20623</a> : 共有作業: LLAP のマップ結合 キャッシュエントリの共有を拡張
バックポート	<a href="#">HIVE-20619</a> : デフォルトで HiveServer2 MultiDelimitSerDe に を含める
バックポート	<a href="#">HIVE-20618</a> : バケット化されていないテーブル では、結合の選択 BucketMapJoin中に選択され る場合があります
バックポート	<a href="#">HIVE-20617</a> : IN 式の定数の型が正しい型にな るように修正
バックポート	<a href="#">HIVE-20612</a> : CBO 用の新しいジョインマルチ キー相関フラグを作成

型	説明
バックポート	<a href="#">HIVE-20603</a> : テーブルロケーションファイルシステムを変更した後にパーティションに挿入すると「Wrong FS」エラーが発生する
バックポート	の ALTER_PARTITION イベントの <a href="#">HIVE-20601</a> : EnvironmentContext null DbNotificationListener
バックポート	<a href="#">HIVE-20583</a> : の kerberos 認証にのみ正規ホスト名を使用する HiveConnection
バックポート	<a href="#">HIVE-20582</a> : hive プロトローキングの hflush を設定可能にする
バックポート	<a href="#">HIVE-20563</a> : ベクトル化: THEN/ELSE 型と結果の型が異なる場合に CASE WHEN 式が失敗する
バックポート	<a href="#">HIVE-20558</a> : hive.hashtable.key.count.adjustment のデフォルトを 0.99 に変更
バックポート	<a href="#">HIVE-20552</a> : スキーマをより速く取得 LogicalPlan する
バックポート	<a href="#">HIVE-20550</a> : beeline を使用して Hive クエリを送信するように WebHCat を切り替える
バックポート	<a href="#">HIVE-20537</a> : CBO と Hive で異なる非相関列を使用した複数列結合推定
バックポート	<a href="#">HIVE-20524</a> : Hive バージョン 2 からバージョン 3 への移行時に、ALTER TABLE VARCHAR から DECIMAL へのスキーマ進化チェックが壊れる

型	説明
バックポート	<a href="#">HIVE-20522</a> : フィールドの null 可能性が原因でアサーションエラーをスロー HiveFilterSetOpTransposeRule することがある
バックポート	<a href="#">HIVE-20521</a> : HS2 doAs=true は hadoop.tmp.dir、MR と S3A ファイルシステムで権限の問題がある
バックポート	<a href="#">HIVE-20515</a> : 結果キャッシュとクエリー時ディレクトリ、結果キャッシュディレクトリを異なるファイルシステムで使用するとクエリ結果が空になる
バックポート	<a href="#">HIVE-20508</a> : Hive は「user@realm」タイプのユーザー名をサポートしない
バックポート	<a href="#">HIVE-20507</a> : Beeline: beeline-site.xml からすべての URI を取得するユーティリティコマンドを追加
バックポート	<a href="#">HIVE-20505</a> : org.openjdk.jmh:jmh-core を 1.21 にアップグレード
バックポート	<a href="#">HIVE-20503</a> : マップ結合選択時にデータ構造を考慮した推定を使用する
バックポート	<a href="#">HIVE-20498</a> : 列統計自動収集の日付型をサポート
バックポート	<a href="#">HIVE-20496</a> : ベクトル化: ベクトル化された PTF IllegalStateException
バックポート	<a href="#">HIVE-20494</a> : HIVE-19440 後に GenericUDFRestrictInformationSchema が壊れる HIVE-19440

型	説明
バックポート	式に INs OptimizedSql が含まれている場合、 <a href="#">HIVE-20477</a> : は表示されません
バックポート	<a href="#">HIVE-20467</a> : リソースプランの作成/削除で IF NOT EXISTS/IF EXISTS を許可する
バックポート	<a href="#">HIVE-20462</a> : ビューが既に存在する場合、「CREATE VIEW IF NOT EXISTS」が失敗する
バックポート	<a href="#">HIVE-20455</a> : security.authorization.PrivilegeSynchronizer.run からのログスプウ
バックポート	<a href="#">HIVE-20439</a> : llap の結合選択時に拡張されたメモリ制限を使用する
バックポート	<a href="#">HIVE-20433</a> : 暗黙の文字列からタイムスタンプへの変換が遅い
バックポート	<a href="#">HIVE-20432</a> : 統計情報推定のために整数型の BETWEEN を IN に書き換える
バックポート	<a href="#">HIVE-20423</a> : NULLS LAST をデフォルトのヌル順序として設定
バックポート	<a href="#">HIVE-20418</a> : LLAP IO が、列が選択されていないクエリの実行インデックスが無効になっている ORC ファイルを正しく処理しない場合がある
バックポート	<a href="#">HIVE-20412</a> : の NPE HiveMetaHook
バックポート	<a href="#">HIVE-20406</a> : 入れ子になった Coalesce で間違った結果が出る
バックポート	<a href="#">HIVE-20399</a> : 完全修飾されていないカスタムテーブルの場所を含む CTAS は MM テーブルで失敗する

型	説明
バックポート	<a href="#">HIVE-20393</a> : 半結合減少: markSemiJoinForDPP の動作に一貫性がない
バックポート	<a href="#">HIVE-20391</a> : HiveAggregateReduceFunctionRule 集計関数を分解するときに誤った戻り型を推測することがある
バックポート	<a href="#">HIVE-20383</a> : hive プロトイベントフックの無効なキュー名と同期の問題。
バックポート	<a href="#">HIVE-20367</a> : ベクトル化: PTF AVG、MAX、MIN、SUM のストリーミングのサポート
バックポート	<a href="#">HIVE-20366</a> : TPC-DS query78 の is null フィルターの統計推定値がずれている
バックポート	<a href="#">HIVE-20364</a> : hive.map.aggr.hash.min.reduction デフォルトをアップデート
バックポート	<a href="#">HIVE-20352</a> : ベクトル化: グループ化関数をサポート
バックポート	<a href="#">HIVE-20347</a> : hive.optimize.sort.dynamic.partition はパーティション化された CTAS と MV で動作する必要がある
バックポート	<a href="#">HIVE-20345</a> : 別の呼び出しからテーブルが削除されると、Drop database がハングアップすることがある
バックポート	<a href="#">HIVE-20343</a> : Hive 3: CTAS は transactional_properties を考慮しない
バックポート	<a href="#">HIVE-20340</a> : タイムスタンプ関数の出力を String として使用する場合、Druid はタイムスタンプから文字列への明示的な CAST を必要とする

型	説明
バックポート	<a href="#">HIVE-20339</a> : ベクトル化: RANK を持つ一部の PTF がベクトル化されない原因となっている不要な制限を解除する
バックポート	<a href="#">HIVE-20337</a> : CachedStore : getPartitionsByExpr がパーティションリストを正しく入力していません
バックポート	<a href="#">HIVE-20336</a> : マテリアライズドビューのマスキングポリシーとフィルタリングポリシー
バックポート	<a href="#">HIVE-20326</a> : NO RELY の代わりに RELY をデフォルトとして制約を作成
バックポート	<a href="#">HIVE-20321</a> : ベクトル化: メモリサイズを 1 コールに減ら VectorHashKeyWrapper して <1 CacheLine
バックポート	<a href="#">HIVE-20320</a> : hive.optimize.remove.sql_count_check フラグをオンにする
バックポート	<a href="#">HIVE-20315</a> : ベクトル化: NULL/誤った結果に関する問題をさらに修正し、不要なキャスト/変換を回避
バックポート	<a href="#">HIVE-20314</a> : マテリアライズドビューの書き換えにパーティションプルーニングを含める
バックポート	<a href="#">HIVE-20312</a> : 矢印クライアント BufferAllocator が独自のを使用できるようにする LlapOutputFormatService
バックポート	<a href="#">HIVE-20302</a> : LLAP: IO でのベクトル化されていない実行では、ROW_ID を含む仮想列が無視される
バックポート	<a href="#">HIVE-20300</a> : VectorFileSinkArrowOperator

型	説明
バックポート	<a href="#">HIVE-20299</a> : LLAP 署名者ユニットテストにおける競合の可能性
バックポート	<a href="#">HIVE-20296</a> : より高度なコンテキストから抽出 HivePointLookupOptimizerRule できるように改善
バックポート	<a href="#">HIVE-20294</a> : ベクトル化: COALESCE/ELT における NULL/間違った結果の問題を修正
バックポート	<a href="#">HIVE-20292</a> : 主制約が定義されている tpcds query93 の結合順序が正しくない
バックポート	<a href="#">HIVE-20290</a> : 中にバッファが割り当てられない ArrowColumnarBatchSerDe ように初期化が遅れる GetSplits
バックポート	<a href="#">HIVE-20281</a> : 「オペレータキャッシュの内容と実際の計画が異なる SharedWorkOptimizer」失敗
バックポート	<a href="#">HIVE-20277</a> : ベクトル化: BOOLEAN を返すケース式は FILTER ではサポートされていない
バックポート	<a href="#">HIVE-20267</a> : WebUI を拡張してログレベルを動的に設定するフォームを含める
バックポート	<a href="#">HIVE-20263</a> : HiveReduceExpressionsWithStatsRule 変数の Typo
バックポート	<a href="#">HIVE-20260</a> : 別の列のフィルターによって行数が変更された場合、列の NDV をスケールしてはならない

型	説明
バックポート	<a href="#">HIVE-20252</a> : Semijoin 削減: 小さいテーブル側の上流にマップ結合があると、セミ結合ブランチによるサイクルが検出されないままになることがある。
バックポート	<a href="#">HIVE-20245</a> : ベクトル化: BETWEEN/IN の NULL/間違った結果の問題を修正
バックポート	<a href="#">HIVE-20241</a> : CTAS ステートメントのパーティショニング仕様をサポート
バックポート	<a href="#">HIVE-20240</a> : Semijoin 削減: ローカル変数を使用して外部テーブル条件をチェック
バックポート	<a href="#">HIVE-20226</a> : リクエスト maxEvents がテーブルの max_rows を超えた場合、HMS getNextNotification は例外をスローする
バックポート	<a href="#">HIVE-20225</a> : Teradata バイナリ形式 SerDe をサポート
バックポート	<a href="#">HIVE-20213</a> : Calcite を 1.17.0 にアップグレード
バックポート	<a href="#">HIVE-20212</a> : http モードの Hiveserver2 がメトリック default.General.open_connections を誤って出力する
バックポート	<a href="#">HIVE-20210</a> : パーティション以外の列のフィルターと変換が最小限 MapReduce である場合、Simple Fetch オプティマイザはにつながる必要がある
バックポート	<a href="#">HIVE-20209</a> : repl ダンプでの最初の試行でメタストア接続が失敗する



型	説明
バックポート	<a href="#">HIVE-20207</a> : ベクトル化: フィルター/比較での NULL/間違っ た結果の問題を修正
バックポート	<a href="#">HIVE-20204</a> : IN 時の型変換
バックポート	<a href="#">HIVE-20203</a> : Arrow が を SerDe リークする DirectByteBuffer
バックポート	<a href="#">HIVE-20197</a> : ベクトル化: DECIMAL_64 テストの追加、日付/間隔/タイムスタンプ算術の追加、および GROUP BY 集計機能の追加
バックポート	<a href="#">HIVE-20193</a> : 説明プランの JSON に cboInfo が ない
バックポート	<a href="#">HIVE-20192</a> : メタストアが埋め込まれた HS2 が JDO PersistenceManager オブジェクトを リークしている
バックポート	<a href="#">HIVE-20183</a> : ソーステーブルに空のバケットが 含まれている場合、バケット化されたテーブル から挿入するとデータが失われる可能性がある
バックポート	<a href="#">HIVE-20177</a> : ベクトル化: GroupBy ストリーミングモードでの割り当てを減らす KeyWrapper
バックポート	<a href="#">HIVE-20174</a> : ベクトル化: GROUP BY 集計関数 における NULL/誤った結果の問題を修正
バックポート	<a href="#">HIVE-20172</a> : リモートメタストアへの接続中に GSS 例外で StatsUpdater 失敗
バックポート	<a href="#">HIVE-20153</a> : Hive 2 以上で UDF をカウントして合計するとメモリ消費量が増える

型	説明
バックポート	<a href="#">HIVE-20152</a> : repl ダンプが失敗したときにデータベースの状態をリセットし、テーブルの名前を変更できるようにする
バックポート	<a href="#">HIVE-20149</a> : TestHiveCli 失敗/タイムアウト
バックポート	<a href="#">HIVE-20130</a> : 情報スキーマシンクロナイザーのロギングの改善
バックポート	<a href="#">HIVE-20129</a> : orc テーブルについてポジションベースのスキーマ進化に戻す
バックポート	<a href="#">HIVE-20118</a> : SessionStateUserAuthenticat or。getGroupNames
バックポート	<a href="#">HIVE-20116</a> : TezTask は親ロガーを使用しています
バックポート	<a href="#">HIVE-20115</a> : Acid テーブルは分析にフッタースキャンを使うべきではない
バックポート	<a href="#">HIVE-20103</a> : WM: 少なくとも 1 つを使用している場合のみ DAG カウンターを集約する
バックポート	<a href="#">HIVE-20101</a> : BloomKFilter: ローカルの byte[] 配列を使用することは完全に避ける
バックポート	<a href="#">HIVE-20100</a> : OpTraits : 不一致が検出されるとオプトレイトの選択が停止する
バックポート	<a href="#">HIVE-20098</a> : 統計: 日付列のパーティション統計を取得する場合の NPE
バックポート	<a href="#">HIVE-20095</a> : jdbc 外部テーブルに計算をプッシュする機能を修正
バックポート	<a href="#">HIVE-20093</a> : LlapOutputFomatService: Netty ArrowBuf で会計に使用する

型	説明
バックポート	<a href="#">HIVE-20090</a> : semijoin 削減フィルターの作成を拡張して新しいオポチュニティを発見できるようにする
バックポート	<a href="#">HIVE-20088</a> : Beeline コンフィグのロケーションパスが正しくアセンブルされていない
バックポート	<a href="#">HIVE-20082</a> : HiveDecimal 文字列変換で 10 進数が正しくフォーマットされない
バックポート	<a href="#">HIVE-20069</a> : DPP と Semijoin の最適化の場合の再最適化を修正
バックポート	<a href="#">HIVE-20051</a> : 一時テーブルの認証をスキップする
バックポート	<a href="#">HIVE-20044</a> : Arrow Serde は文字値を埋め、空の文字列を正しく処理する必要がある
バックポート	<a href="#">HIVE-20028</a> : メタストアクライアントのキャッシュ設定が正しく使用されていない
バックポート	<a href="#">HIVE-20025</a> : によって作成されたイベントファイルのクリーンアップ HiveProtoLoggingHook
バックポート	<a href="#">HIVE-20020</a> : Hive contrib jar はライブラリにあってはいけない
バックポート	<a href="#">HIVE-20013</a> : to_date 関数の日付型への暗黙的なキャストを追加
バックポート	<a href="#">HIVE-20011</a> : プロトロギングフックの追加モードから離れる
バックポート	<a href="#">HIVE-20005</a> : acid_table_stats、acid_no_buckets など - ブランチでのクエリ結果の変更

型	説明
バックポート	<a href="#">HIVE-20004</a> : ConvertDecimal64 によって間違ったスケールが使用されToDecimal、誤った結果になる
バックポート	<a href="#">HIVE-19995</a> : acid テーブルの行トラフィックを集計する
バックポート	<a href="#">HIVE-19993</a> : 列名としても表示されるテーブルエイリアスは使用できない
バックポート	<a href="#">HIVE-19992</a> : ベクトル化: HIVE-19951 へのフォローオン: データ型変換が暗黙的でない場合にのみ ORC のエンコードされた LLAP I/O を無効にするための SchemaEvolution.is OnlyImplicitConversion への呼び出しを追加
バックポート	<a href="#">HIVE-19989</a> : メタストアが HADOOP2 メトリクスに間違ったアプリケーション名を使用する
バックポート	<a href="#">HIVE-19981</a> : HiveStrictManagedMigration ユーティリティによって外部テーブルに変換されたマネージドテーブルは、テーブルが削除されたときにデータを削除するように設定する必要があります
バックポート	<a href="#">HIVE-19967</a> : SMB Join: PTFOperator ala GBY Op に Optraits が必要
バックポート	<a href="#">HIVE-19935</a> : Hive WM セッションが強制終了された: LLAP タスクカウントを更新できなかった
バックポート	<a href="#">HIVE-19924</a> : Repl Load によって実行される distcp ジョブにタグを付ける

型	説明
バックポート	<a href="#">HIVE-19891</a> : カスタムパーティションディレクトリを含む外部テーブルに挿入すると、データが失われる可能性がある
バックポート	<a href="#">HIVE-19850</a> : Tez の動的パーティションプルーニングにより、「No work found for tablescan」というエラーが発生する
バックポート	<a href="#">HIVE-19806</a> : qtests の出力をソートしてテスト結果が不安定にならないようにする
バックポート	<a href="#">HIVE-19770</a> : select に同じ列が複数あるクエリの CBO の Support
バックポート	<a href="#">HIVE-19769</a> : データベースとテーブル名専用のオブジェクトを作成する
バックポート	<a href="#">HIVE-19765</a> : Parquet 固有のテストを に追加する BlobstoreCliDriver
バックポート	<a href="#">HIVE-19759</a> : フラッキーテスト: TestRpc#estServerPort
バックポート	<a href="#">HIVE-19711</a> : Hive スキーマツールのリファクタリング
バックポート	<a href="#">HIVE-19701</a> : getDelegationTokenFromMetaStore 同期する必要はありません
バックポート	<a href="#">HIVE-19694</a> : マテリアライズドビューの作成ステートメントは、MV の SQL ステートメントを実行する前に MV 名の競合をチェックする必要があります。
バックポート	<a href="#">HIVE-19674</a> : Group by 10 進定数が Druid テーブルにプッシュダウンする

型	説明
バックポート	<a href="#">HIVE-19668</a> : 重複する org.antlr.runtime によってヒープの 30% CommonToken 以上が浪費されている。のと重複する文字列
バックポート	<a href="#">HIVE-19663</a> : LLAP IO レポート生成のリファクタリング
バックポート	<a href="#">HIVE-19661</a> : Hive UDF を Re2J 正規表現エンジンを使用するように切り替える
バックポート	<a href="#">HIVE-19628</a> : LLAP testSigning で NPE が発生する可能性がある
バックポート	<a href="#">HIVE-19568</a> : アクティブ/パッシブ HS2 HA: パッシブ HS2 インスタンスへの直接接続を許可しない
バックポート	<a href="#">HIVE-19564</a> : ベクトル化: 算術での NULL/誤った結果の問題を修正
バックポート	<a href="#">HIVE-19552</a> : TestMiniDruidKafkaCliDriver #druidkafkamini_basic.q を有効にする
バックポート	<a href="#">HIVE-19432</a> : hive のデータベースとテーブルが多すぎると速度 GetTablesOperation が遅すぎる
バックポート	<a href="#">HIVE-19360</a> : CBO: QueryPlan オブジェクトに「optimizedSQL」を追加する
バックポート	<a href="#">HIVE-19326</a> : 統計の自動収集: UNION クエリ中の集計が正しくない
バックポート	<a href="#">HIVE-19313</a> : TestJdbcWithDBTokenStoreNoDoAs テストが失敗しています

型	説明
バックポート	<a href="#">HIVE-19285</a> : のサブクラスにログを追加する MetaDataOperation
バックポート	<a href="#">HIVE-19235</a> : Minimr テストのゴールデンファイルを更新
バックポート	<a href="#">HIVE-19104</a> : 再試行でテストを開始する場合、 インスタンス MetaStore は独立している必要があります
バックポート	<a href="#">HIVE-18986</a> : テーブルに多数の列が含まれている場合、 テーブル名を変更すると dataNucleus で java.lang.StackOverflowError in が実行されます
バックポート	<a href="#">HIVE-18920</a> : CBO: 最初のクエリの前に Janino プロバイダーを初期化する
バックポート	<a href="#">HIVE-18873</a> : で MR の述語プッシュダウンをサイレントにスキップ HiveInputFormat すると、 ストレージハンドラーが誤った結果を生成する可能性があります
バックポート	<a href="#">HIVE-18871</a> : hive.aux.jars.path を hdfs:// に設定したために hive on tez 実行エラーが発生する
バックポート	<a href="#">HIVE-18725</a> : 列参照が間違っている場合のサブクエリのエラー処理を改善
バックポート	<a href="#">HIVE-18696</a> : パーティションフォルダは、 の場合、 HiveMetaStore.add_partitions_core メソッドで適切にクリーンアップされないことがあります。

型	説明
バックポート	<a href="#">HIVE-18453</a> : ACID: ACID ORC と Parquet のサポートを統一するために「CREATE TRANSACTIONAL TABLE」構文を追加
バックポート	<a href="#">HIVE-18201</a> : sq_count_chec の XPROD_EDGE を無効にする
バックポート	<a href="#">HIVE-18140</a> : 基本統計が混在していると、パーティションテーブルの統計が間違ってしまうことがある
バックポート	<a href="#">HIVE-17921</a> : LLAP で構造体を使って集計すると間違った結果になる
バックポート	<a href="#">HIVE-17896</a> : TopNKey: ベクトル化可能なスタンドアロンの TopNKey 演算子を作成する
バックポート	transactionalListeners.notifyEvent が失敗した場合の <a href="#">HIVE-17840</a> : HiveMetaStore eats 例外
バックポート	<a href="#">HIVE-17043</a> : 後で参照されない場合、一意でない列を group by キーから削除する
バックポート	<a href="#">HIVE-17040</a> : FK リレーションシップが存在する場合の Join 削除
バックポート	<a href="#">HIVE-16839</a> : 同じパーティションを同時に変更したときの openTransaction/commitTransaction の呼び出しが不均衡になる
バックポート	<a href="#">HIVE-16100</a> : 動的ソートパーティション 옵ティマイザが兄弟演算子を失う
バックポート	<a href="#">HIVE-15956</a> : 多数のパーティションを削除する StackOverflowError 場合



型	説明
バックポート	<a href="#">HIVE-15177</a> : Kerberos 認証タイプが fromSubject に設定されていて、プリンシパルに _HOST が含まれていると、hive による認証が失敗する
バックポート	<a href="#">HIVE-14898</a> : HS2 は空の認証ヘッダーエラーのコールスタックをログに記録すべきでない
バックポート	<a href="#">HIVE-14493</a> : マテリアライズドビューのパーティショニングサポート
バックポート	<a href="#">HIVE-14431</a> : COALESCE を CASE として認識する
バックポート	<a href="#">HIVE-13457</a> : モニタリング情報用の HS2 REST API エンドポイントを作成する
バックポート	<a href="#">HIVE-12342</a> : hive.optimize.index.filter のデフォルト値を true に設定
バックポート	<a href="#">HIVE-10296</a> : hive がメタストアでマルチ結合クエリを実行するとキャスト例外が発生する
バックポート	<a href="#">HIVE-6980</a> : ダイレクト SQL を使用してテーブルをドロップする

## Amazon EMR 6.6.0 - Hive 設定の変更点

- OSS 変更 [HIVE-20703](#) の一環として、動的パーティションをソートするプロパティ `hive.optimize.sort.dynamic.partition` が `hive.optimize.sort.dynamic.partition.threshold` に置き換えられました。

`hive.optimize.sort.dynamic.partition.threshold` 構成には次のような潜在的な値があります。

値	説明
0 (デフォルト)	ORC ファイルを使用する際に、動的パーティションのソートの最適化をコストベースで決定できるようにします。INSERT クエリで許可されるライターの最大数は、(エグゼキューター/コンテナメモリ) * (orc が使用するメモリの割合) を単一のライターが使用する最大メモリ (ストライプサイズ) で割って計算されます。
-1	動的パーティションを完全にソートする最適化を無効にします。
1	動的パーティションのグローバルソートを有効にします。これにより、リデューサー内のパーティション値ごとに 1 つのレコードライターのみが開いたままになるため、リデューサーにかかるメモリ負荷が軽減されます。
2 (またはそれより大きい整数)	指定された整数をライターの最大数のしきい値として使用するよう Hive に指示します。

## Amazon EMR 6.6.0 - Hive の既知の問題

- join と同じ列にウィンドウ関数があるクエリは、[HIVE-25278](#) で報告されているように無効な変換を引き起こし、誤った結果やクエリ失敗の原因となることがあります。回避策として、このようなクエリのクエリレベルで CBO を無効にすることができます。詳細については、AWS サポートにお問い合わせください。
- Amazon EMR 6.6.0 には Hive ソフトウェアバージョン 3.1.2 が含まれています。Hive 3.1.2 では、テキストファイルにヘッダーとフッターが含まれている場合にテキストファイルを分割する機能が導入されています ([HIVE-21924](#))。Apache Tez App Master は各ファイルを読み取り、データ範囲内のオフセットポイントを決定します。クエリで大量の小さなテキストファイルが読み込まれる場合、これらの動作が組み合わさるとパフォーマンスに悪影響を及ぼす可能性があります。回避策として、CombineHiveInputFormat を使用し、以下のプロパティを設定して最大分割サイズを調整してください。

```
SET hive.tez.input.format=org.apache.hadoop.hive.q1.io.CombineHiveInputFormat;  
SET mapreduce.input.fileinputformat.split.maxsize=16777216;
```

- Amazon EMR 6.6.0 から 6.9.x では、動的パーティションと ORDER BY 句または SORT BY 句を使用した INSERT クエリには常に 2 つのリデューサーがあります。この問題は、OSS が変更された [HIVE-20703](#) が原因です。これにより、動的ソートパーティションの最適化がコストベースの決定下に置かれます。ワークロードで動的パーティションのソートが不要な場合は、`hive.optimize.sort.dynamic.partition.threshold` プロパティを -1 に設定して新機能を無効にし、リデューサーの数を正しく計算することをお勧めします。この問題は、[HIVE-22269](#) の一部として OSS Hive で修正され、Amazon EMR 6.10.0 で修正されています。

# Hudi

[Apache Hudi](#) は、増分データの処理とデータパイプラインの開発をシンプルにするオープンソースのデータ管理フレームワークであり、レコードレベルの挿入、更新、アップサート、および削除機能を備えています。アップサートとは、既存のデータセットにレコードがまだ存在しない場合は挿入し、存在する場合は更新する機能を指します。Amazon S3 でのデータのレイアウト方法を効率的に管理することで、Hudi はデータをほぼリアルタイムで取り込んで更新できます。Hudi は、データセットに対して実行されたアクションのメタデータをきめ細かく管理して、アクションをアトミックで整合性のある状態にするのに役立ちます。

Hudi は、[Apache Spark](#)、[Apache Hive](#)、および [Presto](#) と統合されています。Amazon EMR リリースバージョン 6.1.0 以降では、Hudi は [Trino \(PrestoSQL\)](#) とも統合されています。

Amazon EMR リリースバージョン 5.28.0 以降では、EMR は Spark、Hive、Presto、または Flink のインストール時に、デフォルトで Hudi コンポーネントをインストールします。Spark または Hudi DeltaStreamer ユーティリティを使用して、Hudi データセットを作成または更新できます。Hive、Spark、Presto、または Flink を使用して、Hudi データセットをインタラクティブにクエリしたり、増分プルを使用してデータ処理パイプラインを構築したりできます。増分プルとは、2 つのアクション間で変更されたデータのみをプルする機能を指します。

これらの機能により、Hudi は以下のユースケースに適しています。

- 特定のデータの挿入および更新イベントを必要とするセンサーやその他のモノのインターネット (IoT) デバイスからのストリーミングデータを操作する。
- ユーザーがデータの使用方法について忘れてたり、同意を変更したりする可能性のあるアプリケーションのデータプライバシー規制を遵守する。
- 経時的に変更をデータセットに適用できる [変更データキャプチャ \(CDC\) システム](#) を実装する。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Hudi のバージョンと、Amazon EMR で Hudi と共にインストールされるコンポーネントを示しています。

このリリースで Hudi と共にインストールされるコンポーネントのバージョンについては、[「リリース 7.1.0 コンポーネントバージョン」](#) を参照してください。

## emr-7.1.0 の Hudi バージョン情報


Amazon EMR リリースラベル	Hudi バージョン	Hudi でインストールされるコンポーネント
emr-7.1.0	Hudi 0.14.1-amzn-0	Not available.

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Hudi のバージョンと、Amazon EMR で Hudi と共にインストールされるコンポーネントを示しています。

このリリースで Hudi と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の Hudi バージョン情報

Amazon EMR リリースラベル	Hudi バージョン	Hudi でインストールされるコンポーネント
emr-6.15.0	Hudi 0.14.0-amzn-0	Not available.

 Note

Amazon EMR リリース 6.8.0 には [Apache Hudi](#) 0.11.1 が付属していますが、Amazon EMR 6.8.0 クラスタは Hudi 0.12.0 のオープンソース hudi-spark3.3-bundle\_2.12 と互換性があります。

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Hudi のバージョンと、Amazon EMR で Hudi と共にインストールされるコンポーネントを示しています。

このリリースで Hudi と共にインストールされるコンポーネントのバージョンについては、「[リリース 5.36.2 コンポーネントバージョン](#)」を参照してください。

## emr-5.36.2 の Hudi バージョン情報

Amazon EMR リリースラベル	Hudi バージョン	Hudi でインストールされるコンポーネント
emr-5.36.2	Hudi 0.10.1-amzn-1	Not available.

## トピック

- [Hudi の仕組み](#)
- [Amazon EMR で Hudi を使用する際の考慮事項と制限事項](#)
- [Hudi がインストールされたクラスターを作成する](#)
- [Hudi データセットを操作する](#)
- [Hudi CLI を使用する](#)
- [Hudi リリース履歴](#)

## Hudi の仕組み

Amazon EMR で Hudi を使用する場合、Spark データソース API または Hudi DeltaStreamer ユーティリティを使用してデータセットにデータを書き込むことができます。Hudi は、*basepath* の下で、従来の Hive テーブルに似たパーティション分割されたディレクトリ構造にデータセットを編成します。これらのディレクトリ内のファイルとしてのデータのレイアウト方法の詳細は、選択したデータセットタイプによって異なります。書き込み時コピー (CoW) または読み取り時マージ (MoR) を選択できます。

データセットタイプにかかわらず、データセット内の各パーティションは *basepath* に対する *partitionpath* によって一意に識別されます。各パーティション内で、レコードは複数のデータファイルに分散されます。詳細については、Apache Hudi ドキュメントの「[File management](#)」を参照してください。

Hudi の各アクションには、対応するコミットがあり、インスタントと呼ばれる単調に増加するタイムスタンプで識別されます。Hudi は、データセットに対して実行された一連のすべてのアクションをタイムラインとして保持します。Hudi は、タイムラインを使用して、読み取り側と書き込み側のスナップショット分離を提供し、前の時点にロールバックできるようにします。Hudi が記録するアクションとアクションの状態の詳細については、Apache Hudi ドキュメントの「[タイムライン](#)」を参照してください。

## データセットストレージタイプについて: 書き込み時にコピーおよび読み取り時にマージ

Hudi データセットを作成するときに、データセットが書き込み時コピーであるか読み取り時マージであるかを指定します。

- 書き込み時コピー (CoW) – データは列形式 (Parquet) で保存され、更新ごとに書き込み中にファイルの新しいバージョンが作成されます。CoW はデフォルトのストレージタイプです。
- 読み取り時マージ (MoR) – データは、列形式 (Parquet) 形式と行形式 (Avro) の組み合わせを使用して保存されます。更新は、行形式の差分ファイルに記録され、必要に応じて圧縮されて、新しいバージョンの列形式のファイルが作成されます。

CoW データセットでは、レコードが更新されるたびに、そのレコードを含むファイルが更新された値で書き換えられます。MoR データセットでは、更新があるたびに、Hudi によって変更されたレコードの行のみが書き込まれます。MoR は、読み取りが少なく書き込みまたは変更が多いワークロードに適しています。CoW は、頻繁に変更されないデータの読み取りが多いワークロードに適しています。

Hudi は、データにアクセスするための 3 つの論理ビューを提供します。

- 読み取り最適化ビュー – CoW テーブルからの最新のコミットされたデータセットと、MoR テーブルからの最新の圧縮されたデータセットを提供します。
- 増分ビュー – CoW データセットからの 2 つのアクション間の変更ストリームを、ダウンストリームジョブと ETL (抽出、変換、ロード) ワークフローのフィード用に提供します。
- リアルタイムビュー – 列形式のファイルと行形式のファイルをインラインでマージすることで、MoR テーブルからの最新のコミットされたデータを提供します。

読み取り最適化ビューをクエリすると、クエリからすべての圧縮されたデータが返されますが、このデータには最新の差分コミットは含まれていません。このデータをクエリすると、読み取りパフォーマンスは向上しますが、最新のデータは省かれています。リアルタイムビューをクエリすると、圧縮されたデータが Hudi によって読み取り時に差分コミットにマージされます。最新のデータがクエリに使用可能になりますが、マージのコンピューティングオーバーヘッドにより、クエリのパフォーマンスが低下します。圧縮されたデータまたはリアルタイムデータのいずれかをクエリする機能では、クエリ時にパフォーマンスと柔軟性のどちらを優先させるかを選択できます。

ストレージタイプ間での優先項目の選択の詳細については、Apache Hudi ドキュメントの「[Storage Types & Views](#)」を参照してください。

Hudi は MoR のために、Hive Metastore に 2 つのテーブルを作成します。ひとつは読み込み最適化ビューで使用するための、ユーザーが指定したテーブルです。もうひとつはリアルタイムビューで使用するための、同名テーブルの末尾に `_rt` が付与されたテーブルです。両方のテーブルをクエリできます。

## メタストアへの Hudi データセットの登録

Hudi テーブルを Hive メタストアに登録すると、他のテーブルと同様に、Hive、Spark SQL、または Presto を使用して Hudi テーブルをクエリできます。さらに、AWS Glue データカタログをメタストアとして使用するよう Hive と Spark を設定することで、Hudi を AWS Glue と統合できます。MoR テーブルの場合、Hudi はデータセットをメタストアの 2 つのテーブルとして登録します。それらは、指定した名前のテーブル (読み取り最適化ビュー) と、同じ名前でも `_rt` を付加したテーブル (リアルタイムビュー) です。

Spark を使用して Hudi データセットを作成するときに、Hudi テーブルを Hive メタストアに登録します。そのためには、`HIVE_SYNC_ENABLED_OPT_KEY` オプションを `"true"` に設定し、他の必要なプロパティを指定します。詳細については、「[Hudi データセットを操作する](#)」を参照してください。さらに、`hive_sync_tool` コマンドラインユーティリティを使用して、Hudi データセットを個別にテーブルとしてメタストアに登録できます。

## Amazon EMR で Hudi を使用する際の考慮事項と制限事項

- レコードキーフィールドを null または空にすることはできない - レコードキーフィールドとして指定するフィールドには、null または空の値を指定できません。
- アップサートと挿入時にデフォルトで更新されたスキーマ - Hudi は、入力 DataFrame と既存の Hudi データセット `HoodieRecordPayload` をマージして新しい更新されたデータセットを生成する方法を決定するインターフェイスを提供します。Hudi は、このクラスのデフォルト実装を提供します。これにより `OverwriteWithLatestAvroPayload`、既存のレコードが上書きされ、入力 で指定されたスキーマが更新されます DataFrame。マージおよび部分的な更新を実装するためにこのロジックをカスタマイズするには、`DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY` パラメータを使用して `HoodieRecordPayload` インターフェイスの実装を提供できます。
- 削除にはスキーマが必要 - 削除するときは、レコードキー、パーティションキー、事前結合キーのフィールドを指定する必要があります。他の列は null または空にすることができますが、完全なスキーマが必要です。
- MoR テーブルの制限 — MOR テーブルではセーブポイントはサポートされていません。Spark SQL、Presto、または Hive から、読み取り最適化ビューまたはリアルタイムビュー (`tableName_rt`) を使用して MoR テーブルをクエリできます。読み取り最適化ビューを使用した場合、ベースファイルデータのみが公開され、ベースデータとログデータのマージされたビューは公開されません。
- [Hive]



- テーブルを Hive メタストアに登録する場合、Hudi では、Hive Thrift サーバーがデフォルトのポート 10000 で実行されていると想定されます。このポートをカスタムポートで上書きする場合は、以下の例に示すように HIVE\_URL\_OPT\_KEY オプションを渡します。

```
.option(DataSourceWriteOptions.HIVE_URL_OPT_KEY, "jdbc:hive2://localhost:override-port-number
```

- Spark の timestamp データ型は、Hive の timestamp 型としてではなく、Hive の long データ型として登録されます。
- Presto
  - Presto では、Hudi バージョン 0.6.0 未満の場合、MoR リアルタイムテーブルの読み取りはサポートされていません。
  - Presto は、スナップショットクエリのみをサポートしています。
  - Presto で Hudi データセット列が正しく解釈されるためには、hive.parquet\_use\_column\_names 値を true に設定します。
  - セッションの値を設定するには、Presto シェルで以下のコマンドを実行します。

```
set session hive.parquet_use_column_names=true
```

- クラスターレベルで値を設定するには、以下の例に示すように、presto-connector-hive 設定分類を使用して hive.parquet.use\_column\_names を true に設定します。詳細については、「[アプリケーションの設定](#)」を参照してください。

```
[  
  {  
    "Classification": "presto-connector-hive",  
    "Properties": {  
      "hive.parquet.use-column-names": "true"  
    }  
  }  
]
```

- HBase インデックス
  - Hudi の構築に使用した HBase のバージョンが、EMR リリースガイドに記載されているバージョンと異なる場合があります。Spark セッションに対して正しい依存関係をプルインするには、次のコマンドを実行します。

```
spark-shell \  

```

```
--jars /usr/lib/spark/external/lib/spark-avro.jar,/usr/lib/hudi/cli/lib/*.jar \  
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \  
--conf "spark.sql.hive.convertMetastoreParquet=false"
```

## Hudi がインストールされたクラスターを作成する

Amazon EMR リリースバージョン 5.28.0 以降では、Amazon EMR は Spark、Hive、または Presto のインストール時に、デフォルトで Hudi コンポーネントをインストールします。Amazon EMR で Hudi を使用するには、1 つまたは複数の以下のアプリケーションがインストールされたクラスターを作成します。

- Hadoop
- [Hive]
- Spark
- Presto
- Flink

クラスターは、AWS Management Console、AWS CLI または Amazon EMR API を使用して作成できます。

を使用して Hudi でクラスターを作成するには AWS Management Console

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Software Configuration] (ソフトウェア設定) の [Release] (リリース) で emr-5.28.0 以降を選択し、[Hadoop]、[Hive]、[Spark]、[Presto]、[Tez] と、クラスターに必要なその他のアプリケーションを選択します。
4. アプリケーションに必要な他のオプションを設定し、[Next (次へ)] を選択します。
5. 必要に応じて、[ハードウェア] と [General cluster settings (全般クラスター設定)] のオプションを設定します。
6. [Security Options (セキュリティオプション)] で、マスターノードのコマンドラインへの SSH による接続に使用できる [EC2 キーペア] を選択することをお勧めします。これにより、このガイ

ドで説明されている Spark シェルコマンド、Hive CLI コマンド、Hudi CLI コマンドを実行できません。

7. 必要に応じて他のセキュリティオプションを選択し、[クラスターの作成] を選択します。

## Hudi データセットを操作する

Hudi は、Spark を介した Hudi データセットのデータの挿入、更新、削除をサポートしています。詳細については、Apache Hudi ドキュメントの「[Writing Hudi Tables](#)」を参照してください。

以下の例では、インタラクティブな Spark シェルを起動し、Spark submit を使用するか、Amazon EMR Notebooks を使用して、Amazon EMR で Hudi を操作する方法を示しています。Hudi DeltaStreamer ユーティリティやその他のツールを使用してデータセットに書き込むこともできます。このセクション全体を通して、例では、デフォルトの hadoop ユーザーとして、SSH を使用してマスターノードに接続しながら、Spark シェルを使用してデータセットを操作する方法を示しています。

### Amazon EMR 6.7 以降を使用して Spark シェルを起動する

Amazon EMR 6.7.0 以降を使用して、spark-shell、spark-submit、または spark-sql を実行している場合は、次のコマンドを渡します。

#### Note

Amazon EMR 6.7.0 は [Apache Hudi 0.11.0-amzn-0](#) を使用していますが、これは以前の Hudi バージョンよりも大幅に改善されています。詳細については、「[Apache Hudi 0.11.0 Migration Guide](#)」を参照してください。このタブの例にはこれらの変更が反映されています。

プライマリノードで Spark シェルを開くには

1. SSH を使用してプライマリノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してプライマリノードに接続する](#)」を参照してください。
2. 以下のコマンドを入力して、Spark シェルを起動します。PySpark シェルを使用するには、`spark-shell` を `pyspark` に置き換えます。

```
spark-shell --jars /usr/lib/hudi/hudi-spark-bundle.jar \
```

```
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \  
--conf  
"spark.sql.catalog.spark_catalog=org.apache.spark.sql.hudi.catalog.HoodieCatalog"  
\  
--conf "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
```

## Amazon EMR 6.6 以前を使用して Spark シェルを起動する

Amazon EMR 6.6.x 以前を使用して、`spark-shell`、`spark-submit`、または `spark-sql` を実行している場合は、次のコマンドを渡します。

### Note

- Amazon EMR 6.2 および 5.31 以降 (Hudi 0.6.x 以降) では、`spark-avro.jar` を設定から省略できます。
- Amazon EMR 6.5 および 5.35 以降 (Hudi 0.9.x 以降) では、`spark.sql.hive.convertMetastoreParquet=false` を設定から省略できます。
- Amazon EMR 6.6 および 5.36 以降 (Hudi 0.10.x 以降) には、「[Version: 0.10.0 Spark Guide](#)」に記載されている `HoodieSparkSessionExtension` 設定が含まれている必要があります。

```
--conf  
"spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"  
\  

```

プライマリノードで Spark シェルを開くには

1. SSH を使用してプライマリノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してプライマリノードに接続する](#)」を参照してください。
2. 以下のコマンドを入力して、Spark シェルを起動します。PySpark シェルを使用するには、`spark-shell` を `pyspark` に置き換えます。

```
spark-shell \  
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \  
--conf "spark.sql.hive.convertMetastoreParquet=false" \  

```

```
--jars /usr/lib/hudi/hudi-spark-bundle.jar,/usr/lib/spark/external/lib/spark-avro.jar
```

## Amazon EMR 6.7 以降を使用する Amazon EMR Notebooks で Hudi を使用する

Amazon EMR Notebooks で Hudi を使用するには、まずローカルファイルシステムから Hudi jar ファイルをノートブッククラスターのマスターノードの HDFS にコピーする必要があります。次に、ノートブックエディタで、Hudi を使用するように EMR notebook を設定します。

Amazon EMR Notebooks で Hudi を使用するには

1. Amazon EMR Notebooks 用のクラスターを作成して起動します。詳細については、「Amazon EMR 管理ガイド」の「[ノートブックの Amazon EMR クラスターの作成](#)」を参照してください。
2. SSH を使用してクラスターのマスターノードに接続し、以下の例に示すようにローカルファイルシステムから HDFS に jar ファイルをコピーします。この例では、ファイルの管理をわかりやすくするために、HDFS にディレクトリを作成しています。必要に応じて、HDFS で独自のターゲットを選択できます。

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-spark-bundle.jar
```

3. ノートブックエディタを開き、以下の例のコードを入力して実行します。

```
%%configure
{ "conf": {
    "spark.jars": "hdfs:///apps/hudi/lib/hudi-spark-bundle.jar",
    "spark.serializer": "org.apache.spark.serializer.KryoSerializer",
    "spark.sql.catalog.spark_catalog":
"org.apache.spark.sql.hudi.catalog.HoodieCatalog",

    "spark.sql.extensions": "org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
}}
```

## Amazon EMR 6.6 以前を使用する Amazon EMR Notebooks で Hudi を使用する

Amazon EMR Notebooks で Hudi を使用するには、まずローカルファイルシステムから Hudi jar ファイルをノートブッククラスターのマスターノードの HDFS にコピーする必要があります。次に、ノートブックエディタで、Hudi を使用するように EMR notebook を設定します。

Amazon EMR Notebooks で Hudi を使用するには

1. Amazon EMR Notebooks 用のクラスターを作成して起動します。詳細については、「Amazon EMR 管理ガイド」の「[ノートブックの Amazon EMR クラスターの作成](#)」を参照してください。
2. SSH を使用してクラスターのマスターノードに接続し、以下の例に示すようにローカルファイルシステムから HDFS に jar ファイルをコピーします。この例では、ファイルの管理をわかりやすくするために、HDFS にディレクトリを作成しています。必要に応じて、HDFS で独自のターゲットを選択できます。

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-spark-bundle.jar
```

```
hdfs dfs -copyFromLocal /usr/lib/spark/external/lib/spark-avro.jar /apps/hudi/lib/spark-avro.jar
```

3. ノートブックエディタを開き、以下の例のコードを入力して実行します。

```
{ "conf": {  
    "spark.jars": "hdfs:///apps/hudi/lib/hudi-spark-bundle.jar,hdfs:///apps/hudi/lib/spark-avro.jar",  
    "spark.serializer": "org.apache.spark.serializer.KryoSerializer",  
    "spark.sql.hive.convertMetastoreParquet": "false"  
}}
```

## Hudi の Spark セッションを初期化する

Scala を使用する場合、Spark セッションで次のクラスをインポートする必要があります。この操作は、Spark セッションごとに 1 回行う必要があります。

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._
import org.apache.hudi.DataSourceWriteOptions
import org.apache.hudi.DataSourceReadOptions
import org.apache.hudi.config.HoodieWriteConfig
import org.apache.hudi.hive.MultiPartKeyValueExtractor
import org.apache.hudi.hive.HiveSyncConfig
import org.apache.hudi.sync.common.HoodieSyncConfig
```

## Hudi データセットに書き込む

次の例は、を作成し DataFrame、Hudi データセットとして書き込む方法を示しています。

### Note

コードサンプルを Spark シェルに貼り付けるには、プロンプトで「:paste」と入力し、例を貼り付けて、CTRL + D を押します。

Hudi データセット DataFrame に を書き込むたびに、 を指定する必要があります DataSourceWriteOptions。これらのオプションの多くは、書き込みオペレーション間で共通することがよくあります。この例では、*hudiOptions* 変数を使用して、後続の例でも使用する共通のオプションを指定しています。

Amazon EMR 6.7 以降で Scala を使用する場合は書き込み

### Note

Amazon EMR 6.7.0 は [Apache Hudi 0.11.0-amzn-0](#) を使用していますが、これは以前の Hudi バージョンよりも大幅に改善されています。詳細については、「[Apache Hudi 0.11.0 Migration Guide](#)」を参照してください。このタブの例にはこれらの変更が反映されていません。

```
// Create a DataFrame
val inputDF = Seq(
  ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
  ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
  ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
```

```

("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TBL_NAME.key -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE.key -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  HoodieSyncConfig.META_SYNC_PARTITION_EXTRACTOR_CLASS.key ->
"org.apache.hudi.hive.MultiPartKeysValueExtractor",
  HoodieSyncConfig.META_SYNC_ENABLED.key -> "true",
  HiveSyncConfig.HIVE_SYNC_MODE.key -> "hms",
  HoodieSyncConfig.META_SYNC_TABLE_NAME.key -> "tableName",
  HoodieSyncConfig.META_SYNC_PARTITION_FIELDS.key -> "creation_date"
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
  .format("hudi")
  .options(hudiOptions)
  .option(DataSourceWriteOptions.OPERATION_OPT_KEY,"insert")
  .mode(SaveMode.Overwrite)
  .save("s3://DOC-EXAMPLE-BUCKET/myhudidataset/"))

```

## Amazon EMR 6.6 以前で Scala を使用する場合の書き込み

```

// Create a DataFrame
val inputDF = Seq(
  ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
  ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
  ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
  ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
  ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
  ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
).toDF("id", "creation_date", "last_update_time")

```



```
//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TABLE_NAME -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE_OPT_KEY -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY ->
  classOf[MultiPartKeysValueExtractor].getName
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
  .format("org.apache.hudi")
  .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
  DataSourceWriteOptions.INSERT_OPERATION_OPT_VAL)
  .options(hudiOptions)
  .mode(SaveMode.Overwrite)
  .save("s3://DOC-EXAMPLE-BUCKET/myhudidataset/"))
```

## を使用した書き込み PySpark

```
# Create a DataFrame
inputDF = spark.createDataFrame(
  [
    ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
    ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
    ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
    ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
    ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
    ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z"),
  ],
  ["id", "creation_date", "last_update_time"]
)

# Specify common DataSourceWriteOptions in the single hudiOptions variable
hudiOptions = {
'hoodie.table.name': 'tableName',
'hoodie.datasource.write.recordkey.field': 'id',
'hoodie.datasource.write.partitionpath.field': 'creation_date',
```

```
'hoodie.datasource.write.precombine.field': 'last_update_time',
'hoodie.datasource.hive_sync.enable': 'true',
'hoodie.datasource.hive_sync.table': 'tableName',
'hoodie.datasource.hive_sync.partition_fields': 'creation_date',
'hoodie.datasource.hive_sync.partition_extractor_class':
  'org.apache.hudi.hive.MultiPartKeyValueExtractor'
}

# Write a DataFrame as a Hudi dataset
inputDF.write \
  .format('org.apache.hudi') \
  .option('hoodie.datasource.write.operation', 'insert') \
  .options(**hudiOptions) \
  .mode('overwrite') \
  .save('s3://DOC-EXAMPLE-BUCKET/myhudidataset/')
```

### Note

コードの例や通知には、Hudi ではなく「hoodie」と表示されている場合があります。Hudi コードベースは、古い「hoodie」の継りが広く使用されています。

## DataSourceWriteOptions Hudi の リファレンス

オプション	説明
TABLE_NAME	データセットを登録するテーブルの名前。
TABLE_TYPE_OPT_KEY	オプション。データセットが "COPY_ON_WRITE" として作成されるか、"MERGE_ON_READ" として作成されるかを指定します。デフォルトは "COPY_ON_WRITE" です。
RECORDKEY_FIELD_OPT_KEY	値が HoodieKey の recordKey コンポーネントとして使用されるレコードキーフィールド。実際の値は、フィールド値に対して <code>.toString()</code> を呼び出すことで取得されます。ネストされたフィールドは、a.b.c など、ドット表記を使用して指定できます。

オプション	説明
PARTITIONPATH_FIELD_OPT_KEY	値が HoodieKey の partitionPath コンポーネントとして使用されるパーティションパスフィールド。実際の値は、フィールド値に対して <code>.toString()</code> を呼び出すことで取得されます。
PRECOMBINE_FIELD_OPT_KEY	実際の書き込みの前に事前結合で使用されるフィールド。2つのレコードのキー値が同じ場合、Hudi は、 <code>Object.compareTo(...)</code> で決定された事前結合フィールドの値が最も大きいレコードを選択します。

以下のオプションは、Hudi データセットテーブルをメタストアに登録するためにのみ必要です。Hudi データセットを Hive メタストアのテーブルとして登録しない場合、これらのオプションは必要ありません。

#### DataSourceWriteOptions Hive の リファレンス

オプション	説明
HIVE_DATABASE_OPT_KEY	同期先の Hive データベース。デフォルトは "default" です。
HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY	パーティションフィールド値を Hive パーティション列に抽出するために使用されるクラス。
HIVE_PARTITION_FIELDS_OPT_KEY	Hive パーティション列を決定するために使用するデータセット内のフィールド。
HIVE_SYNC_ENABLED_OPT_KEY	"true" に設定すると、データセットを Apache Hive メタストアに登録します。デフォルトは "false" です。
HIVE_TABLE_OPT_KEY	必須。同期先の Hive テーブルの名前。例えば "my_hudi_table_cow" です。

オプション	説明
HIVE_USER_OPT_KEY	オプション。同期時に使用する Hive ユーザー名。例えば "hadoop" です。
HIVE_PASS_OPT_KEY	オプション。HIVE_USER_OPT_KEY で指定されたユーザーの Hive パスワード。
HIVE_URL_OPT_KEY	Hive メタストアの URL。

## データのアップサート

次の例は、 を記述してデータをアップサートする方法を示しています DataFrame。以前の挿入の例とは異なり、OPERATION\_OPT\_KEY 値は UPSERT\_OPERATION\_OPT\_VAL に設定されています。また、.mode(SaveMode.Append) は、レコードを付加する必要があることを指示するために指定されています。

Amazon EMR 6.7 以降で Scala を使用する Upsert

### Note

Amazon EMR 6.7.0 は [Apache Hudi 0.11.0-amzn-0](#) を使用していますが、これは以前の Hudi バージョンよりも大幅に改善されています。詳細については、「[Apache Hudi 0.11.0 Migration Guide](#)」を参照してください。このタブの例にはこれらの変更が反映されています。

```
// Create a new DataFrame from the first row of inputDF with a different creation_date value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
  .format("hudi")
  .options(hudiOptions)
  .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "upsert")
  .mode(SaveMode.Append)
  .save("s3://DOC-EXAMPLE-BUCKET/myhuidataset/"))
```

## Amazon EMR 6.6 以前で Scala を使用する Upsert

```
// Create a new DataFrame from the first row of inputDF with a different creation_date value
val updatedDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updatedDF.write
  .format("org.apache.hudi")
  .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
    DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
  .options(hudiOptions)
  .mode(SaveMode.Append)
  .save("s3://DOC-EXAMPLE-BUCKET/myhuidataset/"))
```

## を使用したアップサート PySpark

```
from pyspark.sql.functions import lit

# Create a new DataFrame from the first row of inputDF with a different creation_date value
updateDF = inputDF.limit(1).withColumn('creation_date', lit('new_value'))

updateDF.write \
  .format('org.apache.hudi') \
  .option('hoodie.datasource.write.operation', 'upsert') \
  .options(**hudiOptions) \
  .mode('append') \
  .save('s3://DOC-EXAMPLE-BUCKET/myhuidataset/')
```

## レコードを削除する

レコードをハード削除するには、空のペイロードをアップサートします。この場合、PAYLOAD\_CLASS\_OPT\_KEY オプションで EmptyHoodieRecordPayload クラスを指定しています。この例では DataFrame、アップサートの例で使用されているのと同じ updateDF、を使用して、同じレコードを指定します。

## Amazon EMR 6.7 以降で Scala を使用する場合の削除

### Note

Amazon EMR 6.7.0 は [Apache Hudi 0.11.0-amzn-0](#) を使用していますが、これは以前の Hudi バージョンよりも大幅に改善されています。詳細については、「[Apache Hudi 0.11.0](#)

[Migration Guide](#) を参照してください。このタブの例にはこれらの変更が反映されていません。

```
(updateDF.write
  .format("hudi")
  .options(hudiOptions)
  .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "delete")
  .mode(SaveMode.Append)
  .save("s3://DOC-EXAMPLE-BUCKET/myhuidataset/"))
```

### Amazon EMR 6.6 以前で Scala を使用する場合の削除

```
(updateDF.write
  .format("org.apache.hudi")
  .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
  .option(DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY,
"org.apache.hudi.common.model.EmptyHoodieRecordPayload")
  .mode(SaveMode.Append)
  .save("s3://DOC-EXAMPLE-BUCKET/myhuidataset/"))
```

### を使用して を削除する PySpark

```
updateDF.write \
  .format('org.apache.hudi') \
  .option('hoodie.datasource.write.operation', 'upsert') \
  .option('hoodie.datasource.write.payload.class',
'org.apache.hudi.common.model.EmptyHoodieRecordPayload') \
  .options(**hudiOptions) \
  .mode('append') \
  .save('s3://DOC-EXAMPLE-BUCKET/myhuidataset/')
```

OPERATION\_OPT\_KEY を DELETE\_OPERATION\_OPT\_VAL に設定して、送信するデータセット内のすべてのレコードを削除し、データをハード削除することもできます。ソフト削除の実行方法、および Hudi テーブルに保存されたデータの削除の詳細については、Apache Hudi ドキュメントの「[Deletes](#)」(削除) を参照してください。

## Hudi データセットから読み込む

現時点でのデータを取得するために、Hudi はデフォルトでスナップショットクエリを実行します。以下に、[Hudi データセットに書き込む](#) で S3 に書き込まれたデータセットをクエリする例を示します。`s3://DOC-EXAMPLE-BUCKET/myhudidataset` をテーブルパスに置き換え、パーティションレベルごとにワイルドカードのアスタリスクを追加し、さらにアスタリスクを 1 つ追加します。この例では、パーティションレベルが 1 つあるため、2 つのワイルドカードシンボルを追加しました。

Amazon EMR 6.7 以降で Scala を使用する場合は読み込み

### Note

Amazon EMR 6.7.0 は [Apache Hudi 0.11.0-amzn-0](#) を使用していますが、これは以前の Hudi バージョンよりも大幅に改善されています。詳細については、「[Apache Hudi 0.11.0 Migration Guide](#)」を参照してください。このタブの例にはこれらの変更が反映されています。

```
val snapshotQueryDF = spark.read
  .format("hudi")
  .load(s3://DOC-EXAMPLE-BUCKET/myhudidataset)
  .show()
```

Amazon EMR 6.6 以前で Scala を使用する場合は読み込み

```
(val snapshotQueryDF = spark.read
  .format("org.apache.hudi")
  .load("s3://DOC-EXAMPLE-BUCKET/myhudidataset" + "/*/*"))

snapshotQueryDF.show()
```

を使用した読み取り PySpark

```
snapshotQueryDF = spark.read \
  .format('org.apache.hudi') \
  .load('s3://DOC-EXAMPLE-BUCKET/myhudidataset' + '/*/*')

snapshotQueryDF.show()
```

## 増分クエリ

Hudi で増分クエリを実行して、特定のコミットタイムスタンプ以降に変更されたレコードのストリームを取得することもできます。これを行うには、QUERY\_TYPE\_OPT\_KEY フィールドを QUERY\_TYPE\_INCREMENTAL\_OPT\_VAL に設定します。次に、BEGIN\_INSTANTTIME\_OPT\_KEY の値を追加して、指定された時刻以降に書き込まれたすべてのレコードを取得します。増分クエリでは、変更されたレコードのみが処理されるため、通常、バッチクエリよりも 10 倍効率が高くなります。

増分クエリを実行するときは、スナップショットクエリに使用されるワイルドカードのアスタリスクを付けずに、ルート (ベース) テーブルパスを使用します。

### Note

Presto は増分クエリをサポートしていません。

## Scala を使用した増分クエリ

```
(val incQueryDF = spark.read
  .format("org.apache.hudi")
  .option(DataSourceReadOptions.QUERY_TYPE_OPT_KEY,
    DataSourceReadOptions.QUERY_TYPE_INCREMENTAL_OPT_VAL)
  .option(DataSourceReadOptions.BEGIN_INSTANTTIME_OPT_KEY, <beginInstantTime>)
  .load("s3://DOC-EXAMPLE-BUCKET/myhudidataset" ))

incQueryDF.show()
```

## を使用した増分クエリ PySpark

```
readOptions = {
  'hoodie.datasource.query.type': 'incremental',
  'hoodie.datasource.read.begin.instanttime': <beginInstantTime>,
}

incQueryDF = spark.read \
  .format('org.apache.hudi') \
  .options(**readOptions) \
  .load('s3://DOC-EXAMPLE-BUCKET/myhudidataset')

incQueryDF.show()
```



Hudi データセットからの読み取りの詳細については、Apache Hudi ドキュメントの「[Querying Hudi tables](#)」(Hudiテーブルをクエリ)を参照してください。

## Hudi CLI を使用する

Hudi CLI を使用して Hudi データセットを管理し、コミット、ファイルシステム、統計などに関する情報を表示できます。CLI を使用して、手動で圧縮を実行したり、圧縮をスケジュールしたり、スケジュールされた圧縮をキャンセルしたりすることもできます。詳細については、Apache Hudi ドキュメントの「[Interacting via CLI](#)」を参照してください。

Hudi CLI を起動してデータセットに接続するには

1. SSH を使用してマスターノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してマスターノードに接続する](#)」を参照してください。
2. コマンドラインで「/usr/lib/hudi/cli/bin/hudi-cli.sh」と入力します。

コマンドプロンプトが `hudi->` に変わります。

3. 以下のコードを入力して、データセットに接続します。`s3://DOC-EXAMPLE-BUCKET/myhudidataset` は、処理するデータセットへのパスに置き換えます。使用する値は、前の例で設定した値と同じです。

```
connect --path s3://DOC-EXAMPLE-BUCKET/myhudidataset
```

以下の例に示すように、コマンドプロンプトが変わって、接続されているデータセットが表示されます。

```
hudi:myhudidataset->
```

## Hudi リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Hudi のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

## Hudi バージョン情報

Amazon EMR リリースラベル	Hudi バージョン	Hudi でインストールされるコンポーネント
emr-5.36.2	0.10.1-amzn-1	Not available.
emr-7.1.0	0.14.1-amzn-0	Not available.
emr-7.0.0	0.14.0-amzn-1	Not available.
emr-6.15.0	0.14.0-amzn-0	Not available.
emr-6.14.0	0.13.1-amzn-2	Not available.
emr-6.13.0	0.13.1-amzn-1	Not available.
emr-6.12.0	0.13.1-amzn-0	Not available.
emr-6.11.1	0.13.0-amzn-0	Not available.
emr-6.11.0	0.13.0-amzn-0	Not available.
emr-6.10.1	0.12.2-amzn-0	Not available.
emr-6.10.0	0.12.2-amzn-0	Not available.
emr-6.9.1	0.12.1-amzn-0	Not available.
emr-6.9.0	0.12.1-amzn-0	Not available.
emr-6.8.1	0.11.1-amzn-0	Not available.
emr-6.8.0	0.11.1-amzn-0	Not available.
emr-6.7.0	0.11.0-amzn-0	Not available.
emr-5.36.1	0.10.1-amzn-1	Not available.
emr-5.36.0	0.10.1-amzn-1	Not available.
emr-6.6.0	0.10.1-amzn-0	Not available.

Amazon EMR リリースラベル	Hudi バージョン	Hudi でインストールされるコンポーネント
emr-5.35.0	0.9.0-amzn-2	Not available.
emr-6.5.0	0.9.0-amzn-1	Not available.
emr-6.4.0	0.8.0-amzn-0	Not available.
emr-6.3.1	0.7.0-amzn-0	Not available.
emr-6.3.0	0.7.0-amzn-0	Not available.
emr-6.2.1	0.6.0-amzn-1	Not available.
emr-6.2.0	0.6.0-amzn-1	Not available.
emr-6.1.1	0.5.2-incubating-amzn-2	Not available.
emr-6.1.0	0.5.2-incubating-amzn-2	Not available.
emr-6.0.1	0.5.0-incubating-amzn-1	Not available.
emr-6.0.0	0.5.0-incubating-amzn-1	Not available.
emr-5.34.0	0.9.0-amzn-0	Not available.
emr-5.33.1	0.7.0-amzn-1	Not available.
emr-5.33.0	0.7.0-amzn-1	Not available.
emr-5.32.1	0.6.0-amzn-0	Not available.
emr-5.32.0	0.6.0-amzn-0	Not available.
emr-5.31.1	0.6.0-amzn-0	Not available.
emr-5.31.0	0.6.0-amzn-0	Not available.
emr-5.30.2	0.5.2-incubating	Not available.
emr-5.30.1	0.5.2-incubating	Not available.

Amazon EMR リリースラベル	Hudi バージョン	Hudi でインストールされるコンポーネント
emr-5.30.0	0.5.2-incubating	Not available.
emr-5.29.0	0.5.0-incubating	Not available.
emr-5.28.1	0.5.0-incubating	Not available.
emr-5.28.0	0.5.0-incubating	Not available.

# Hue

Hue (Hadoop User Experience) は、Amazon EMR と Apache Hadoop で使用するオープンソースでウェブベースのグラフィカルユーザーインターフェイスです。Hue は、複数の異なる Hadoop エコシステムプロジェクトを設定可能なインターフェイスにまとめます。Amazon EMR は、Amazon EMR 上に Hue に固有のカスタマイズを追加しました。Hue はクラスターで実行されるアプリケーションのフロントエンドとして動作し、より使い慣れた、またはユーザーフレンドリなインターフェイスを使用してアプリケーションを操作することができます。Hive Editor、Pig Editor などの Hue のアプリケーションでは、各アプリケーションのそれぞれのシェルでインタラクティブにスクリプトを実行するために、クラスターにログインする必要がありません。クラスターが起動したら、Hue や同様のインターフェイスを使用してアプリケーションを完全に操作できます。Hue の詳細については、<http://gethue.com> を参照してください。

Hue は、Amazon EMR コンソールを使用してクラスターを起動するときに、デフォルトでインストールされます。Hue をインストールしない選択をするには、クラスターの起動時に Amazon EMR コンソールの [Advanced options] (詳細オプション) を使用する、または、`--applications` から `create-cluster` を使用するとき AWS CLI オプションを明示的に指定して Hue を省略します。

## トピック

- [Hue バージョン情報](#)
- [Amazon EMR で Hue のサポートされている機能とサポート外の機能](#)
- [考慮事項](#)
- [Hue Web ユーザーインターフェイスへの接続](#)
- [Amazon RDS でリモートデータベースと Hue を使用する](#)
- [Hue の詳細設定](#)
- [Hue リリース履歴](#)

## Hue バージョン情報

### 7.1.0 の Hue バージョン

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Hue のバージョンと、Amazon EMR で Hue と共にインストールされるコンポーネントを示しています。

このリリースで Hue と共にインストールされるコンポーネントのバージョンについては、[「リリース 7.1.0 コンポーネントバージョン」](#) を参照してください。

## emr-7.1.0 の Hue バージョン情報

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-7.1.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

## 6.15.0 用 Hue バージョン

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Hue のバージョンと、Amazon EMR で Hue と共にインストールされるコンポーネントを示しています。

このリリースで Hue と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の Hue バージョン情報

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.15.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
		timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

### 5.36.2 の Hue バージョン

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Hue のバージョンと、Amazon EMR で Hue と共にインストールされるコンポーネントを示しています。

このリリースで Hue と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

#### emr-5.36.2 の Hue バージョン情報

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.36.2	Hue 4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

## Amazon EMR で Hue のサポートされている機能とサポート外の機能

Amazon EMR 7.0.0 以降では、Hue には Python 3.9 以降が必要です。転送中の暗号化を使用する場合は、2048 ビットの暗号キーを使用する必要があります。

- Amazon S3 および Hadoop File System (HDFS) Browser
  - 適切なアクセス権限があれば、エフェメラル HDFS ストレージとアカウントに属する S3 バケット間でデータを参照し、移動することができます。
  - デフォルトでは、Hue のスーパーユーザーは、Amazon EMR の IAM ロールがアクセスを許可されているすべてのファイルにアクセスできます。新しく作成されたユーザーには、Amazon S3 ファイルブラウザへのアクセス許可は自動的に付与されません。グループに対して `filebrowser.s3_access` アクセス許可を有効にする必要があります。
- Hive - データに対してインタラクティブなクエリを実行します。これは、プログラムまたはバッチ処理されたクエリのプロトタイプを作成する場合にも役立ちます。
- Pig - データに対してスクリプトを実行、またはインタラクティブコマンドを発行します。
- Oozie - Oozie のワークフローを作成し監視します。
- Metastore Manager - Hive メタストアのコンテンツの表示と操作 (インポート/作成、削除など) を行います。
- ジョブブラウザ - 送信した Hadoop ジョブのステータスを表示します。
- ユーザー管理 - Hue アカウントを管理し、LDAP ユーザーを Hue と統合します。
- AWS サンプル — Hue のアプリケーションを使用してさまざまな AWS のサービスのサンプルデータを処理する ready-to-run 「」 例がいくつかあります。Hue にログインすると、Hue Home アプリケーションに移動します。ここでは、サンプルがプレインストールされています。
- Livy サーバーは Amazon EMR バージョン 5.9.0 以降のみでサポートされます。
- Spark の Hue ノートブックを使用するには、Livy と Spark で Hue をインストールする必要があります。
- Trino/Presto - Hue は Trino データサーバーと Presto データサーバーへの接続をサポートしています。Amazon EMR 7.0.0 以降では、この接続を促進するために `trino-python-client` と `presto-python-client` コネクタを使用しています。7.0.0 より前のリリースでは、JDBC ゲートウェイ経由で JDBC コネクタを使用します。
- Hue ダッシュボードはサポートされていません。
- PostgreSQL はサポートされていません。

## 考慮事項

Amazon EMR で [Hue](#) を使用する場合は、次の制限事項を考慮してください。



## 大きな Hue メタデータテーブルでのパフォーマンス

Hue メタデータデータベースが大きすぎると、パフォーマンスが低下する可能性があります。テーブルのサイズを確認するには、まず SSH を使用して Amazon EC2 クラスター上の Amazon EMR のプライマリノードに接続し、コマンドを実行して MySQL CLI `sudo mysql -u root` を起動します。テーブルのサイズを取得するには、クエリを実行します `SELECT COUNT(*) FROM hue.<table_name>`。<table\_name> の仕組みについては、以下を参照してください。

- デスクトップドキュメント
- デスクトップドキュメント2
- oozie\_job
- beeswax\_savedquery
- beeswax\_session
- beeswax\_queryhistory

そのクエリを実行すると 100,000 を超えるカウントが返された場合は、次のクリーンアップコマンドを実行して古いレコードを削除する必要があります。

```
cd /opt/cloudera/parcels/CDH/lib/hue # Hue home directory ./build/env/bin/hue
desktop_document_cleanup
```

データベースのクリーンアップの詳細については、[Hue ドキュメントの「リファレンスページ」](#)を参照してください。

Hue はテーブルを自動的にクリーンアップしませんが、Amazon EMR リリース 5.12.0 以降では、テーブル内の古いドキュメントを削除する方法が用意されています。次のシェルスクリプトを作成し、メタデータデータベースにドキュメントを保持する最大日数を表す整数パラメータを持つ Amazon EMR クラスターのステップとして実行します。

```
#!/bin/bash
if grep isMaster /mnt/var/lib/info/instance.json | grep false;
then
    echo "This is not the primary node; do nothing, exiting"
    exit 0
fi
while [ ! -f /usr/lib/hue/desktop/core/src/desktop/management/commands/
desktop_document_cleanup.py ]
do
```

```
sleep 1
done
sudo systemctl stop hue.service
sudo sed -i 's+ LOG.warn+ # LOG.warn+g' /usr/lib/hue/desktop/core/src/desktop/
management/commands/desktop_document_cleanup.py
sudo /usr/lib/hue/build/env/bin/hue desktop_document_cleanup --keep-days $1
sudo systemctl start hue.service
```

## Hue バージョン間の非互換性

複数の Hue 対応クラスターで同じ Hue メタデータデータベースを使用している場合は、これらのクラスターで同じバージョンの Hue を実行することをお勧めします。Hue の異なるバージョンでは、Hue メタデータデータベースのスキーマが異なる場合があります。異なるバージョンに同じデータベースを使用すると、Hue のインストールが失敗する可能性があります。例えば、4.10.0 と 4.11.0 がインストールされている 2 つのクラスターに同じデータベースを使用しようとする、ユーザーが 4.10.0 クラスターにログインしようとする、ログインエラーが発生する可能性があります。

Amazon EMR リリースラベルとそれに対応するインストール済みバージョンの Hue のリストについては、[「Hue リリース履歴」](#)を参照してください。

## Hue Web ユーザーインターフェイスへの接続

Hue Web ユーザーインターフェイスへの接続は、クラスターのマスターノードにホストされている任意の HTTP インターフェイスへの接続と同じです。次の手順は、Hue ユーザーインターフェイスにアクセスする方法を示しています。詳細については、「Amazon EMR 管理ガイド」の「[EMR クラスターでホストされるウェブインターフェイスの表示](#)」を参照してください。

Hue ウェブユーザーインターフェイスを表示するには

1. 「Amazon EMR 管理ガイド」の「[ダイナミックポートフォワーディングを使用してマスターノードへの SSH トンネルをセットアップする](#)」の手順に従います。
2. ブラウザに以下のアドレスを入力して [Hue] ウェブインターフェイスを開きます。http://*master public DNS*:8888 ここで、##### *DNS* は、クラスターマスターノードのパブリック DNS 名です (例: ec2-11-22-333-44.compute-1.amazonaws.com)。
3. Hue のログイン画面で、管理者として初めてログインする場合、ユーザー名とパスワードを入力して Hue のスーパーアカウントを作成し、[アカウントの作成] を選択します。それ以外の場合は、ユーザー名とパスワードを入力し、[Create account (アカウントの作成)] を選択するか、管理者から提供された認証情報を入力します。

## Amazon RDS でリモートデータベースと Hue を使用する

デフォルトでは、Hue のユーザー情報とクエリ履歴は、マスターノード上のローカルの MySQL データベースに保存されます。または、Amazon S3 に保存されている設定と、Amazon Relational Database Service(Amazon RDS) の MySQL データベースを使用して、Hue が有効な 1 つ以上のクラスターを作成することができます。こうすることで、Amazon EMR クラスターを実行し続けなくても、Hue によって作成されたユーザー情報とクエリ履歴を永続化できます。設定ファイルは、Amazon S3 サーバーサイド暗号化機能を使って保存することをお勧めします。

最初に、Hue のリモートデータベースを作成します。

外部の MySQL データベースを作成するには

1. Amazon RDS コンソール (<https://console.aws.amazon.com/rds/>) を開きます。
2. [Launch a DB Instance (DB インスタンスの作成)] をクリックします。
3. MySQL を選択し、[Select (選択)] をクリックします。
4. [Multi-AZ Deployment and Provisioned IOPS Storage (マルチ AZ 配置とプロビジョンド IOPS ストレージ)] はデフォルトのままにして、[Next (次へ)] をクリックします。
5. [Instance Specifications (インスタンスの仕様)] はデフォルトのままにして、[Settings (設定)] を指定し、[Next (次へ)] をクリックします。
6. [Configure Advanced Settings] ページで、適切なセキュリティグループとデータベース名を選択します。使用するセキュリティグループは、少なくともクラスターのマスターノードからのポート 3306 に対するインGRESS TCP アクセスを許可する必要があります。この時点でまだクラスターを作成していない場合、ポート 3306 への接続をすべてのホストに許可し、クラスターを起動した後にセキュリティグループを調整することができます。[Launch DB Instance (DB インスタンスの起動)] をクリックします。
7. RDS ダッシュボードで、[Instances (インスタンス)] を選択し、先ほど作成したインスタンスを選択します。データベースが使用できる時、dbname、username、password、および RDS instance hostname をメモします。クラスターの作成および設定時に、この情報を使用します。

を使用してクラスターを起動するときに Hue の外部 MySQL データベースを指定するには AWS CLI

を使用してクラスターを起動するときに Hue の外部 MySQL データベースを指定するには AWS CLI、RDS インスタンスの作成時にメモした情報を使用して、設定オブジェクトhue.iniで を設定します。

**Note**

外部にある同一のデータベースを使用するが、各クラスターはクエリ履歴とユーザー情報を共有する複数のクラスターを作成できます。

- を使用して AWS CLI、Hue がインストールされたクラスターを作成し、作成した外部データベースを使用し、データベースプロパティを指定する Hue の設定分類を持つ設定ファイルを参照します。次の例では、データベース設定を指定する Amazon S3 の設定ファイル `myConfig.json` を参照して、Hue がインストールされた状態でクラスターを作成します。

**Note**

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --release-label emr-7.1.0 --applications Name=Hue Name=Spark
Name=Hive \  
--instance-type m5.xlarge --instance-count 3 \  
--configurations https://s3.amazonaws.com/mybucket/myfolder/myConfig.json --use-  
default-roles
```

`myConfig.json` ファイルのコンテンツの例を以下に示しま

す。`dbname`、`username`、`password`、および `RDS instance hostname` は、先ほど RDS ダッシュボードでメモした値に置き換えます。

```
[{  
  "Classification": "hue-ini",  
  "Properties": {},  
  "Configurations": [  
    {  
      "Classification": "desktop",  
      "Properties": {},  
      "Configurations": [  
        {  
          "Classification": "database",  
          "Properties": {
```

```
        "name": "dbname",
        "user": "username",
        "password": "password",
        "host": "RDS instance hostname",
        "port": "3306",
        "engine": "mysql"
    },
    "Configurations": []
}
]
}
]
}]
```

## トラブルシューティング

### Amazon RDS フェイルオーバーの場合

Hue データベースインスタンスが応答しなくなったり、フェイルオーバー処理の途中であるために、クエリの実行中に遅延が発生する可能性があります。この問題に関する情報とガイドラインを以下に示します。

- Amazon RDS コンソールにログインすると、フェイルオーバーイベントを検索できます。たとえば、フェイルオーバーが処理中または発生したことが表示されたら、"Multi-AZ instance failover started" や "Multi-AZ instance failover completed" などのイベントを探します。
- RDS インスタンスがフェイルオーバーを完了するのに、約 30 秒かかります。
- Hue でクエリに対する longer-than-normal レスポンスが表示されている場合は、クエリを再実行してみてください。

## Hue の詳細設定

このセクションでは、次のトピックについて説明します。

### トピック

- [LDAP ユーザーに関して Hue を設定する](#)

## LDAP ユーザーに関して Hue を設定する

LDAP との統合により、ユーザーは LDAP のディレクトリに保存されている既存の認証情報を使用して Hue にログインすることができます。Hue を LDAP と統合する場合、Hue で個別にユーザー情報を管理する必要はありません。以下に示す情報では、Microsoft Active Directory との Hue の統合を示していますが、その設定オプションはどの LDAP のディレクトリでも類似しています。

LDAP 認証では、最初にサーバーをバインドして接続を確立します。次に、確立された接続は、LDAP ユーザー情報を検索するために以降のクエリで使用されます。Active Directory サーバーで匿名接続が許可されていない限り、接続はバインド識別子名とパスワードで確立される必要があります。バインド識別子名 (DN) は、`bind_dn` 設定で定義されます。バインドパスワードは `bind_password` 設定で定義されます。Hue には、LDAP リクエストをバインドするために、検索バインドと直接バインドという 2 つの方法があります。Amazon EMR で Hue を使用するための推奨の方法は、検索バインドです。

検索バインドが Active Directory で使用される場合、Hue はベース識別子名 (DN) から取得される必要がある属性を見つけるために、ユーザー名属性 (`user_name_attr` config で定義されます) を使用します。検索バインドは、完全な DN が Hue のユーザーに対して不明な場合に役立ちます。

たとえば、共通名 (CN) を使用するように `user_name_attr` config を設定している場合があります。その場合は、Active Directory サーバーは、ログイン時に提供される Hue ユーザー名を使用して、ベース識別子名から開始して一致する共通名をディレクトリツリーで検索します。Hue ユーザーの共通名が見つかった場合、ユーザーの識別子名がサーバーによって返されます。次に、Hue はバインド操作を実行して、ユーザーの認証に使用される識別子名を構築します。

### Note

検索バインドでは、ベース識別子名で開始してすべてのディレクトリのサブツリーでユーザー名を検索します。Hue の LDAP 設定で指定されたベース識別子名はユーザー名の最も近い親である必要があります。それ以外の場合、LDAP 認証のパフォーマンスが低下する可能性があります。

直接バインドが Active Directory で使用されている場合、認証には正確な `nt_domain` または `ldap_username_pattern`、が使用される必要があります。直接バインドが使用されている場合、`nt` ドメイン (`nt_domain` 設定で定義されます) 属性が定義されていると、`<login username>@nt_domain` 形式を使用してユーザー識別名テンプレートが作成されます。このテンプレートは、ベース識別子名で開始してすべてのディレクトリのサブツリーを検索するために使用されます。`nt` ドメインが設定されていない場合、Hue はユーザーの正確な識別子名

パターン ( `ldap_username_pattern` 設定で定義されます ) を検索します。このインスタンスで、サーバーは、ベース識別子名で開始してすべてのディレクトリサブツリーで、一致する `ldap_username_pattern` 値を検索します。

AWS CLIを使用して Hue の LDAP プロパティを伴うクラスターを起動するには

- `hue-ini` の LDAP プロパティを指定するには、Hue がインストールされたクラスターを作成し、LDAP の設定プロパティのある `json` ファイルを参照します。サンプルコマンドを以下に示します。Amazon S3 に格納されている設定ファイル `myConfig.json` を参照します。

```
aws emr create-cluster --release-label emr-7.1.0 --applications Name=Hue Name=Spark
Name=Hive \
--instance-type m5.xlarge --instance-count 3 --configurations https://
s3.amazonaws.com/mybucket/myfolder/myConfig.json.
```

`myConfig.json` のコンテンツの例を以下に示します。

```
[
  {
    "Classification": "hue-ini",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "desktop",
        "Properties": {},
        "Configurations": [
          {
            "Classification": "ldap",
            "Properties": {},
            "Configurations": [
              {
                "Classification": "ldap_servers",
                "Properties": {},
                "Configurations": [
                  {
                    "Classification": "yourcompany",
                    "Properties": {
                      "base_dn":
"DC=yourcompany,DC=hue,DC=com",
                      "ldap_url": "ldap://ldapurl",
                      "search_bind_authentication": "true",
```

```

        "bind_dn":
"CN=hue,CN=users,DC=yourcompany,DC=hue,DC=com",
        "bind_password": "password"
    },
    "Configurations": []
}
]
}
]
},
{
    "Classification": "auth",
    "Properties": {
        "backend": "desktop.auth.backend.LdapBackend"
    }
}
]
}
]
}
]
]

```

### Note

Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

Hue で LDAP 設定を表示するには

1. Amazon EMR クラスターのマスターノードに対するアクティブな VPN 接続または SSH トンネルがあることを確認します。次に、ブラウザで *master-public-dns*:8888 と入力して Hue ウェブインターフェイスを開きます。
2. Hue の管理者認証情報を使用してログインします。[Did you know? (ご存知でしたか?)] ウィンドウが開いた場合は、[Got it, prof! (分かりました!)] をクリックして閉じます。



3. ツールバーの [Hue] アイコンをクリックします。
4. [Hue の概要] ページで、[設定] を選択します。
5. [Configuration Sections and Variables (設定セクションと変数)] セクションで、[Desktop (デスクトップ)] をクリックします。
6. [ldap] セクションにスクロールして、設定を確認します。

## Hue リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Hue のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Hue バージョン情報

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.36.2	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.1.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
		nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.0.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.15.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.14.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.13.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.12.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.10.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.10.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.9.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.9.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.8.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.8.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.7.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server



Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.36.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.6.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.35.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.3.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.3.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.2.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.2.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.1.1	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.1.0	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-6.0.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.0.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.34.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.33.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server



Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.33.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.32.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.32.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.31.1	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.31.0	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.30.2	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.30.1	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.30.0	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.29.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.28.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.28.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.27.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.27.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.26.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.25.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.24.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server



Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.24.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.23.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.23.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.22.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.21.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.21.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.20.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.20.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.19.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.19.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.18.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.18.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.17.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.17.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.17.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server



Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.16.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.16.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.15.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.15.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.14.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.14.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.13.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.13.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.3	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.12.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.12.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.4	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.11.3	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.2	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server



Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.11.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.10.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.10.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.9.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.9.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.8.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.8.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.7.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.7.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.6.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.6.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.4	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.5.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server



Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.5.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.4.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.4.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.3.2	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.2.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.2.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.1.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.1.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.0.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-5.0.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.9.6	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-4.9.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-4.9.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server



Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-4.8.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-4.8.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-4.7.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.7.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-4.6.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.6.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.5.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-4.4.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.3.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.2.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Amazon EMR リリースラベル	Hue バージョン	Hue でインストールされるコンポーネント
emr-4.1.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

# Iceberg

[Apache Iceberg](#) は、Amazon Simple Storage Service (Amazon S3) の大規模データセット用のオープンテーブル形式です。大きなテーブルに対する高速のクエリパフォーマンス、アトミックコミット、同時書き込み、および SQL 互換テーブルの進化を提供します。Amazon EMR 6.5.0 以降では、Iceberg テーブル形式の Amazon EMR クラスターで Apache Spark 3 を使用できます。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Iceberg のバージョンと、Amazon EMR で Iceberg と共にインストールされるコンポーネントを示しています。

このリリースで Iceberg と共にインストールされるコンポーネントのバージョンについては、[「リリース 7.1.0 コンポーネントバージョン」](#) を参照してください。

## emr-7.1.0 の Iceberg バージョン情報

Amazon EMR リリースラベル	Iceberg バージョン	Iceberg でインストールされるコンポーネント
emr-7.1.0	Iceberg 1.4.3-amzn-0	Not available.

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Iceberg のバージョンと、Amazon EMR で Iceberg と共にインストールされるコンポーネントを示しています。

このリリースで Iceberg と共にインストールされるコンポーネントのバージョンについては、[「リリース 6.15.0 コンポーネントバージョン」](#) を参照してください。

## emr-6.15.0 の Iceberg バージョン情報

Amazon EMR リリースラベル	Iceberg バージョン	Iceberg でインストールされるコンポーネント
emr-6.15.0	Iceberg 1.4.0-amzn-0	Not available.

## トピック

- [Iceberg の仕組み](#)
- [Iceberg がインストールされたクラスターを使用する](#)

- [Amazon EMR で Iceberg を使用する際の考慮事項と制限事項](#)
- [Iceberg リリース履歴](#)

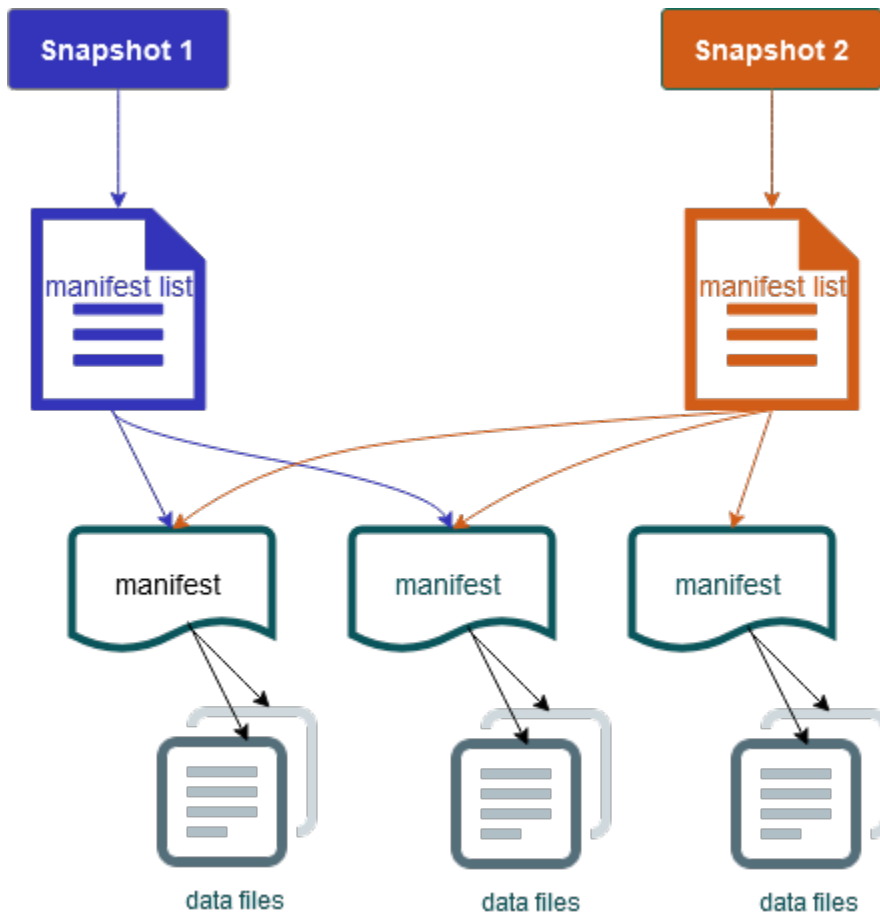
## Iceberg の仕組み

Iceberg は個々のデータファイルをディレクトリではなくテーブルで追跡します。これにより、作成者はデータファイルを作成するだけですみます (ファイルの移動や変更はありません)。また、作成者は明示的なコミットでのみテーブルにファイルを追加できます。テーブルの状態はメタデータファイルで管理されます。テーブル状態を変更すると、古いメタデータをアトミックに置き換える新しいメタデータファイルが作成されます。テーブルメタデータファイルでは、テーブルスキーマ、パーティショニング設定、その他のプロパティが追跡されます。

テーブルコンテンツのスナップショットも含まれています。各スナップショットは、ある時点におけるテーブル内のデータファイルの完全なセットです。スナップショットはメタデータファイルにリスト化されますが、スナップショット内のファイルは別のマニフェストファイルに保存されます。あるテーブルメタデータファイルから次のテーブルメタデータファイルへのアトミックな移行により、スナップショットが分離されます。読み込みでは、テーブルメタデータをロードした時点の最新のスナップショットを使用します。また、更新して新しいメタデータの場所を取得するまで、変更の影響を受けません。スナップショット内のデータファイルは 1 つまたは複数のマニフェストファイルに保存されます。マニフェストファイルには、テーブル内の各データファイル、そのパーティションデータとメトリクスが 1 行ずつ格納されます。スナップショットは、マニフェスト内のすべてのファイルをまとめたものです。また、マニフェストファイルはスナップショット間で共有できるため、変更頻度の低いメタデータを書き換える必要がありません。

### Iceberg スナップショットの図





Iceberg には以下の機能があります。

- Amazon S3 データレイクの ACID トランザクションとタイムトラベルをサポートします。
- コミットの再試行には、[オプティミスティック同時実行](#)のパフォーマンス上の利点が活かされず。
- ファイルレベルの競合解決により、高い同時実行性が実現します。
- メタデータの列ごとの最小/最大統計を使用すると、ファイルをスキップできるため、選択的クエリのパフォーマンスが向上します。
- テーブルを柔軟なパーティションレイアウトに整理でき、パーティションの進化によりパーティションスキームを更新できます。これにより、クエリやデータの量を物理ディレクトリに依存せずに変更できます。
- [スキーマ進化](#)と適用をサポートします。
- Iceberg テーブルは冪等性シンクと再生可能なソースとして機能します。これにより、1 回限りのパイプラインによるストリーミングとバッチサポートが可能になります。冪等性シンクは、過去に

成功した書き込み操作を追跡します。そのため、シンクは障害が発生した場合はデータを再度要求し、複数回送信された場合はデータをドロップできます。

- テーブル進化、操作履歴、各コミットの統計など、履歴と系統を表示します。
- データ形式 (Parquet、ORC、Avro) と分析エンジン (Spark、Trino、PrestoDB、Flink、Hive) を選択して、既存のデータセットから移行します。

## Iceberg がインストールされたクラスターを使用する

このセクションでは、Spark、Trino、Flink、および Hive で Iceberg を使用方法について説明します。

### Spark で Iceberg クラスターを使用する

Amazon EMR バージョン 6.5.0 以降では、ブートストラップアクションを追加しなくても Spark クラスターで Iceberg を使用できます。Amazon EMR バージョン 6.4.0 以前の場合、ブートストラップアクションを使用して必要なすべての依存関係を事前インストールできます。

このチュートリアルでは、を使用して Amazon EMR Spark クラスターで Iceberg AWS CLI を操作します。コンソールを使用して Iceberg がインストールされたクラスターを作成するには、「[Build an Apache Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue](#)」の手順に従ってください。

### Iceberg クラスターの作成

、、AWS CLI または Amazon EMR API を使用して AWS Management Console、Iceberg がインストールされたクラスターを作成できます。このチュートリアルでは、を使用して Amazon EMR クラスターで Iceberg AWS CLI を操作します。コンソールを使用して Iceberg がインストールされたクラスターを作成するには、「[Build an Apache Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue](#)」の手順に従ってください。

で Amazon EMR で Iceberg を使用するには AWS CLI、まず次のステップでクラスターを作成します。を使用して Iceberg 分類を指定する方法については AWS CLI、[クラスターの作成 AWS CLI 時にを使用して設定を指定する](#)「」または「」を参照してください[クラスター作成時に Java SDK を使用して設定を指定する](#)。

1. 以下のコンテンツを含む `configurations.json` ファイルを作成します。

```
[{
```

```
"Classification":"iceberg-defaults",
"Properties":{"iceberg.enabled":"true"}
}]
```

- 次に、以下の設定でクラスターを作成します。この例の Amazon S3 バケットパスとサブネット ID は、実際の値に置き換えてください。

```
aws emr create-cluster --release-label emr-6.5.0 \
--applications Name=Spark \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_Spark_Iceberg_Cluster \
--log-uri s3://DOC-EXAMPLE-BUCKET/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes
InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

または、Spark アプリケーションを含む Amazon EMR クラスターを作成し、`/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` ファイルを Spark ジョブの JAR 依存関係として追加することもできます。詳細については、「[Submitting Applications](#)」を参照してください。

この jar を Spark ジョブの依存関係として含めるには、以下の設定プロパティを Spark アプリケーションに追加します。

```
--conf "spark.jars=/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar"
```

Spark ジョブの依存関係の詳細については、Apache Spark ドキュメント「[Running Spark on Kubernetes](#)」の「[Dependency Management](#)」を参照してください。

## Iceberg の Spark セッションを初期化する

以下の例では、インタラクティブな Spark シェルを起動し、Spark submit を使用するか、Amazon EMR Notebooks を使用して、Amazon EMR で Iceberg を操作する方法を示します。

spark-shell

- SSH を使用してマスターノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してマスターノードに接続する](#)」を参照してください。

2. 以下のコマンドを入力して、Spark シェルを起動します。PySpark シェルを使用するには、`spark-shell`に置き換えます `pyspark`。

```
spark-shell \  
--conf  
  "spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"  
 \  
--conf "spark.sql.catalog.dev=org.apache.iceberg.spark.SparkCatalog" \  
--conf "spark.sql.catalog.dev.type=hadoop" \  
--conf "spark.sql.catalog.dev.warehouse=s3://DOC-EXAMPLE-BUCKET/example-prefix/"
```

## spark-submit

1. SSH を使用してマスターノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してマスターノードに接続する](#)」を参照してください。
2. 以下のコマンドを入力して、Iceberg の Spark セッションを起動します。

```
spark-submit \  
--conf  
  "spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"  
 \  
--conf "spark.sql.catalog.dev=org.apache.iceberg.spark.SparkCatalog" \  
--conf "spark.sql.catalog.dev.type=hadoop" \  
--conf "spark.sql.catalog.dev.warehouse=s3://DOC-EXAMPLE-BUCKET/example-prefix/"
```

## EMR Studio notebooks

EMR Studio ノートブックを使用して Spark セッションを初期化するには、次の例のように、Amazon EMR Notebooks で `%%configure` マジックコマンドを使用して Spark セッションを設定します。詳細については、「Amazon EMR 管理ガイド」の「[EMR Notebooks マジックを使用する](#)」を参照してください。

```
%%configure -f  
{  
  "conf":{  
  
    "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions",  
    "spark.sql.catalog.dev":"org.apache.iceberg.spark.SparkCatalog",  
    "spark.sql.catalog.dev.type":"hadoop",
```

```
"spark.sql.catalog.dev.warehouse": "s3://DOC-EXAMPLE-BUCKET/example-prefix/"
}
}
```

## Iceberg テーブルへの書き込み

次の例は、を作成して Iceberg データセットとして DataFrame 書き込む方法を示しています。この例では、デフォルトの Hadoop ユーザーとして SSH を使用してマスターノードに接続しながら、Spark シェルを使用してデータセットを操作する方法を示しています。

### Note

コードサンプルを Spark シェルに貼り付けるには、プロンプトで「:paste」と入力し、例を貼り付けて、[CTRL+D] を押します。

## PySpark

Spark には、Python ベースのシェルである pyspark 用意されており、Python で記述された Spark プログラムのプロトタイプを作成するために使用できます。マスターノードで pyspark を起動します。

```
## Create a DataFrame.
data = spark.createDataFrame([
    ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
    ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
    ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
    ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
],["id", "creation_date", "last_update_time"])

## Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://DOC-EXAMPLE-BUCKET/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

## Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame.
val data = Seq(
  ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
  ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
  ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
  ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
).toDF("id", "creation_date", "last_update_time")

// Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://DOC-EXAMPLE-BUCKET/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

## Iceberg テーブルからの読み込み

### PySpark

```
df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

### Scala

```
val df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

### Spark SQL

```
SELECT * from dev.db.iceberg_table LIMIT 10
```

## AWS Glue Data Catalog を Iceberg テーブルメタストアとして使用するよう Spark プロパティを設定する

AWS Glue Catalog を Iceberg テーブルのメタストアとして使用するには、Spark 設定プロパティを次のように設定します。

```
spark-submit \  
  --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \  
  --conf spark.sql.catalog.my_catalog.warehouse=s3://<bucket>/<prefix> \  
  --conf spark.sql.catalog.my_catalog.catalog-  
impl=org.apache.iceberg.aws.glue.GlueCatalog \  
  --conf spark.sql.catalog.my_catalog.io-impl=org.apache.iceberg.aws.s3.S3FileIO \  
  --conf spark.sql.catalog.my_catalog.lock-  
impl=org.apache.iceberg.aws.dynamodb.DynamoDbLockManager \  
  --conf spark.sql.catalog.my_catalog.lock.table=myGlueLockTable
```

## Trino での Iceberg クラスターの使用

Amazon EMR バージョン 6.6.0 以降では、Iceberg を Trino クラスターで使用できます。

このチュートリアルでは、を使用して Amazon EMR Trino クラスターで Iceberg AWS CLI を操作します。コンソールを使用して Iceberg がインストールされたクラスターを作成するには、「[Build an Apache Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue](#)」の手順に従ってください。

### Iceberg クラスターの作成

で Amazon EMR で Iceberg を使用するには AWS CLI、まず次のステップでクラスターを作成します。を使用して Iceberg 分類を指定する方法については AWS CLI、[クラスターの作成 AWS CLI 時にを使用して設定を指定する](#)「」または「」を参照してください[クラスター作成時に Java SDK を使用して設定を指定する](#)。

1. iceberg.properties ファイルを作成し、選択したカタログの値を設定します。例えば、Hive メタストアをカタログとして使用する場合、ファイルには次の内容が含まれている必要があります。

```
connector.name=iceberg  
hive.metastore.uri=thrift://localhost:9083
```

AWS Glue データカタログをストアとして使用する場合は、ファイルに次のコンテンツが含まれている必要があります。

```
connector.name=iceberg
iceberg.catalog.type=glue
```

2. 次の例のように、Amazon S3 の `iceberg.properties` を `/etc/trino/conf/catalog/iceberg.properties` にコピーするブートストラップアクションを作成します。ブートストラップアクションの詳細については、「[追加のソフトウェアをインストールするためのブートストラップアクションの作成](#)」を参照してください。

```
set -ex
sudo aws s3 cp s3://DOC-EXAMPLE-BUCKET/iceberg.properties /etc/trino/conf/catalog/iceberg.properties
```

3. 以下の設定でクラスターを作成し、この例のブートストラップアクションのスクリプトパスとキー名を自分のものに置き換えます。

```
aws emr create-cluster --release-label emr-6.7.0 \
--applications Name=Trino \
--region us-east-1 \
--name My_Trino_Iceberg_Cluster \
--bootstrap-actions '[{"Path":"s3://DOC-EXAMPLE-BUCKET", "Name":"Add iceberg.properties"}]' \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=c3.4xlarge InstanceGroupType=CORE,InstanceCount=3,InstanceType=c3.4xlarge \
--use-default-roles \
--ec2-attributes KeyName=<key-name>
```

## Iceberg の Trino セッションの初期化

Trino セッションを初期化するには、次のコマンドを実行します。

```
trino-cli --catalog iceberg
```

## Iceberg テーブルへの書き込み

次の SQL コマンドを使用してテーブルを作成し、書き込みます。



```
trino> SHOW SCHEMAS;
trino> CREATE TABLE default.iceberg_table (
    id int,
    data varchar,
    category varchar)
WITH (
    format = 'PARQUET',
    partitioning = ARRAY['category', 'bucket(id, 16)'],
    location = 's3://DOC-EXAMPLE-BUCKET/<prefix>')

trino> INSERT INTO default.iceberg_table VALUES (1,'a','c1'), (2,'b','c2'),
(3,'c','c3');
```

## Iceberg のテーブルからの読み込み

Iceberg テーブルから読み込むには、次のコマンドを実行します。

```
trino> SELECT * from default.iceberg_table;
```

## Flink での Iceberg クラスターの使用

Amazon EMR バージョン 6.9.0 以降では、オープンソースの Iceberg Flink 統合を使用する際に必要なセットアップ手順なしに Flink クラスターで Iceberg を使用できます。

### Iceberg クラスターの作成

Iceberg がインストールされたクラスターは、AWS Management Console、AWS CLI、または Amazon EMR API を使用して作成できます。このチュートリアルでは、を使用して Amazon EMR クラスターで Iceberg AWS CLI を操作します。コンソールを使用して Iceberg がインストールされたクラスターを作成するには、「[Build an Apache Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue](#)」の手順に従ってください。

で Amazon EMR で Iceberg を使用するには AWS CLI、まず次のステップでクラスターを作成します。を使用して Iceberg 分類を指定する方法については AWS CLI、[クラスターの作成 AWS CLI 時にを使用して設定を指定する](#)「」または「」を参照してください[クラスター作成時に Java SDK を使用して設定を指定する](#)。configurations.json というファイルを次の内容で作成します。

```
[{
  "Classification":"iceberg-defaults",
  "Properties":{"iceberg.enabled":"true"}
```

```
}]
```

次に、以下の設定でクラスターを作成し、この例の Amazon S3 バケットパスとサブネット ID を独自の値に置き換えます。

```
aws emr create-cluster --release-label emr-6.9.0 \  
--applications Name=Flink \  
--configurations file://iceberg_configurations.json \  
--region us-east-1 \  
--name My_flink_Iceberg_Cluster \  
--log-uri s3://DOC-EXAMPLE-BUCKET/ \  
--instance-type m5.xlarge \  
--instance-count 2 \  
--service-role EMR_DefaultRole \  
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

または、Flink アプリケーションを含む Amazon EMR 6.9.0 クラスターを作成し、`/usr/share/aws/iceberg/lib/iceberg-flink-runtime.jar` ファイルを Flink ジョブの JAR 依存関係として使用することもできます。

## Flink SQL クライアントの使用

SQL クライアントスクリプトは `/usr/lib/flink/bin` にあります。次のコマンドを使用してスクリプトを実行します。

```
flink-yarn-session -d # starting the Flink YARN Session in detached mode  
./sql-client.sh
```

これにより Flink SQL シェルが起動します。

## Flink の例

### Iceberg テーブルの作成

#### Flink SQL

```
CREATE CATALOG glue_catalog WITH (  
  'type'='iceberg',  
  'warehouse'='<WAREHOUSE>',  
  'catalog-impl'='org.apache.iceberg.aws.glue.GlueCatalog',  
  'io-impl'='org.apache.iceberg.aws.s3.S3FileIO',  
  'lock-impl'='org.apache.iceberg.aws.dynamodb.DynamoDbLockManager',
```

```
'lock.table'='myGlueLockTable'
);

USE CATALOG glue_catalog;

CREATE DATABASE IF NOT EXISTS <DB>;

USE <DB>;

CREATE TABLE IF NOT EXISTS `glue_catalog`.`<DB>`.`sample` (id int, data string);
```

## Table API

```
EnvironmentSettings settings =
    EnvironmentSettings.newInstance().inBatchMode().build();

TableEnvironment tEnv = TableEnvironment.create(settings);

String warehouse = "<WAREHOUSE>";
String db = "<DB>";

tEnv.executeSql(
    "CREATE CATALOG glue_catalog WITH (\n"
    + "  'type'='iceberg',\n"
    + "  'warehouse'='"
    + warehouse
    + "',\n"
    + "  'catalog-impl'='org.apache.iceberg.aws.glue.GlueCatalog',
\n"
    + "  'io-impl'='org.apache.iceberg.aws.s3.S3FileIO'\n"
    + " );");

tEnv.executeSql("USE CATALOG glue_catalog;");
tEnv.executeSql("CREATE DATABASE IF NOT EXISTS " + db + ";");
tEnv.executeSql("USE " + db + ";");
tEnv.executeSql(
    "CREATE TABLE `glue_catalog`.`" + db + "`.`sample` (id bigint, data string);");
```

## Iceberg テーブルへの書き込み

### Flink SQL

```
INSERT INTO `glue_catalog`.`<DB>`.`sample` values (1, 'a'),(2,'b'),(3,'c');
```

## Table API

```
tEnv.executeSql(
    "INSERT INTO `glue_catalog`.``"
    + db
    + "`.`sample` values (1, 'a'),(2,'b'),(3,'c');");
```

## Datastream API

```
final StreamExecutionEnvironment env =
    StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

GenericRowData rowData1 = new GenericRowData(2);
rowData1.setField(0, 1L);
rowData1.setField(1, StringData.fromString("a"));

DataStream<RowData> input = env.fromElements(rowData1);

Map<String, String> props = new HashMap<>();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatalogLoader =
    CatalogLoader.custom(
        "glue",
        props,
        new Configuration(),
        "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
    TableLoader.fromCatalog(glueCatalogLoader, TableIdentifier.of(db, "sample"));

DataStreamSink<Void> dataStreamSink =
    FlinkSink.forRowData(input).tableLoader(tableLoader).append();

env.execute("Datastream Write");
```

## Iceberg テーブルからの読み込み

### Flink SQL

```
SELECT * FROM `glue_catalog`.`<DB>`.`sample`;
```

### Table API

```
Table result = tEnv.sqlQuery("select * from `glue_catalog`.`" + db + "`.`sample`");
```

### Datastream API

```
final StreamExecutionEnvironment env =
    StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

Map<String, String> props = new HashMap<>();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatalogLoader =
    CatalogLoader.custom(
        "glue",
        props,
        new Configuration(),
        "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
    TableLoader.fromCatalog(glueCatalogLoader, TableIdentifier.of(db, "sample"));

DataStream<RowData> batch =

    FlinkSource.forRowData().env(env).tableLoader(tableLoader).streaming(false).build();

batch.print().name("print-sink");
```

## Hive カタログの使用

Flink と Hive の依存関係が「[Hive Metastore と Glue Catalog を使用して Flink を設定する](#)」の説明に従って解決されていることを確認します。

## Flink ジョブの実行

Flink にジョブを送信する方法の 1 つは、ジョブ単位の Flink YARN セッションを使用することです。これを実行するには、次のコマンドを使用します。

```
sudo flink run -m yarn-cluster -p 4 -yjm 1024m -ytm 4096m $JAR_FILE_NAME
```

## Hive での Iceberg クラスターの使用

Amazon EMR リリース 6.9.0 以降では、オープンソースの Iceberg Hive 統合に必要なセットアップ手順を実行しなくても、Hive クラスターで Iceberg を使用できます。Amazon EMR バージョン 6.8.0 以前の場合は、ブートストラップアクションを使用して `iceberg-hive-runtime jar` をインストールし Hive での Iceberg サポートを設定できます。

Amazon EMR 6.9.0 には、[Hive 3.1.3 を Iceberg 0.14.1 と統合する](#)ためのすべての機能が含まれています。また、サポートされている実行エンジンを実行時に自動選択するなどの Amazon EMR の追加機能も含まれています (Amazon EMR on EKS 6.9.0)。

## Iceberg クラスターの作成

、 、 AWS CLI または Amazon EMR API を使用して AWS Management Console、Iceberg がインストールされたクラスターを作成できます。このチュートリアルでは、 を使用して Amazon EMR クラスターで Iceberg AWS CLI を操作します。コンソールを使用して Iceberg がインストールされたクラスターを作成するには、「[Build an Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue](#)」の手順を実行します。

で Amazon EMR で Iceberg を使用するには AWS CLI、まず以下のステップを使用してクラスターを作成します。AWS CLI または Java SDK を使用して Iceberg 分類を指定する方法については、[クラスターの作成 AWS CLI 時に を使用して設定を指定する](#)「」または「」を参照してください[クラスター作成時に Java SDK を使用して設定を指定する](#)。以下の内容で `configurations.json` という名前のファイルを作成します。

```
[{
  "Classification": "iceberg-defaults",
  "Properties": {"iceberg.enabled": "true"}
```

```
}]
```

次に、以下の設定でクラスターを作成し、この例の Amazon S3 バケットパスとサブネット ID を自分のものに置き換えます。

```
aws emr create-cluster --release-label emr-6.9.0 \  
--applications Name=Hive \  
--configurations file://iceberg_configurations.json \  
--region us-east-1 \  
--name My_hive_Iceberg_Cluster \  
--log-uri s3://DOC-EXAMPLE-BUCKET/ \  
--instance-type m5.xlarge \  
--instance-count 2 \  
--service-role EMR_DefaultRole \  
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

Hive Iceberg クラスターは以下のことを行います。

- Iceberg Hive ランタイム jar を Hive に読み込み、Hive エンジンの Iceberg 関連の設定を有効にします。
- Amazon EMR Hive の動的実行エンジン選択を有効にして、ユーザーが Iceberg と互換性のある実行エンジンを設定できないようにします。

#### Note

Hive Iceberg クラスターは現在 AWS Glue データカタログをサポートしていません。デフォルトの Iceberg カタログは HiveCatalog で、これは Hive 環境用に設定されたメタストアに対応しています。カタログ管理の詳細については、Apache Hive [ドキュメントのHCatalogの使用](https://cwiki.apache.org/confluence/display/HIVE)」を参照してください。 <https://cwiki.apache.org/confluence/display/HIVE>

## 機能のサポート

Amazon EMR 6.9.0 は Hive 3.1.3 と Iceberg 0.14.1 をサポートしています。機能のサポートについては Hive 3.1.2 と 3.1.3 の Iceberg 互換機能に限定されています。以下の コマンドがサポートされています。

- Amazon EMR リリース 6.9.0~6.12.x では、libfb303 jar を Hive の auxlib ディレクトリに配置する必要があります。これを行うには、次のコマンドを使用します。

```
sudo /usr/bin/ln -sf /usr/lib/hive/lib/libfb303-*.jar /usr/lib/hive/auxlib/  
libfb303.jar
```

Amazon EMR リリース 6.13 以降では、libfb303 jar は自動的に Hive の auxlib ディレクトリにシンボリックリンクされます。


- テーブルの作成

- 非パーティションテーブル - Hive の外部テーブルは、次のようにストレージハンドラーを指定することで作成できます。

```
CREATE EXTERNAL TABLE x (i int) STORED BY  
'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
```

- パーティションテーブル - Hive の外部パーティションテーブルは次のように作成できます。

```
CREATE EXTERNAL TABLE x (i int) PARTITIONED BY (j int) STORED BY  
'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
```

 Note

ORC/AVRO/PARQUET の STORED AS ファイル形式は Hive 3 ではサポートされていません。デフォルトで、唯一のオプションは Parquet です。

- テーブルの削除 - DROP TABLE コマンドは、次の例のようにテーブルを削除するために使用されます。

```
DROP TABLE [IF EXISTS] table_name [PURGE];
```

- テーブルの読み込み - 次の例のように、SELECT ステートメントを使用して Hive の Iceberg テーブルを読み込むことができます。サポートされている実行エンジンは MR と Tez です。

```
SELECT * FROM table_name
```

Hive の選択構文については、[LanguageManual 「の選択」](#)を参照してください。Hive で Iceberg テーブルを使用する select ステートメントの詳細については、「[Apache Iceberg Select](#)」を参照してください。

- テーブルへの挿入 - HiveQL の INSERT INTO ステートメントは Map Reduce 実行エンジンのみをサポートする Iceberg テーブルで動作します。Amazon EMR Hive は実行時に Iceberg テーブルの



エンジンを選択するため、Amazon EMR ユーザーは実行エンジンを明示的に設定する必要はありません。

- 単一テーブルの insert into - 例:

```
INSERT INTO table_name VALUES ('a', 1);
INSERT INTO table_name SELECT...;
```

- 複数テーブルの insert into - 非アトミック複数テーブル insert into ステートメントがサポートされています。例:

```
FROM source
INSERT INTO table_1 SELECT a, b
INSERT INTO table_2 SELECT c,d;
```

## Amazon EMR で Iceberg を使用する際の考慮事項と制限事項

このセクションでは、Spark、Trino、Flink、および Hive で Iceberg を使用する際の考慮事項と制限事項について説明します。

### Spark で Iceberg を使用するための考慮事項

- Amazon EMR 6.5.0 は、デフォルトでは Amazon EMR on EKS での Iceberg の実行をサポートしていません。Amazon EMR 6.5.0 カスタムイメージが用意されているため、`--jars local:///usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` を `spark-submit` パラメータとして渡して Amazon EMR on EKS に Iceberg テーブルを作成できます。詳細については、「Amazon EMR on EKS 開発ガイド」の「[カスタムイメージを使用して Amazon EMR で Spark ワークロードを送信する](#)」を参照してください。AWS Support に問い合わせることもできます。Amazon EMR 6.6.0 以降、Iceberg は Amazon EMR on EKS でサポートされます。
- AWS Glue を Iceberg のカタログとして使用する場合は、テーブルを作成するデータベースが AWS Glue に存在することを確認してください。などのサービスを使用して AWS Lake Formation いて、カタログをロードできない場合は、コマンドを実行するためのサービスへの適切なアクセス権があることを確認してください。

### Trino で Iceberg を使用するための考慮事項

- Amazon EMR 6.5 では、Iceberg の Trino Iceberg Catalog ネイティブサポートは提供されていません。Trino には Iceberg v0.11 が必要です。そのため Trino 用の Amazon EMR クラスターを Spark

クラスターとは別に起動し、そのクラスターに Iceberg v0.11 をインストールすることをお勧めします。

- AWS Glue を Iceberg のカタログとして使用する場合は、テーブルを作成するデータベースが AWS Glue に存在することを確認してください。などのサービスを使用して AWS Lake Formation いて、カタログをロードできない場合は、コマンドを実行するためのサービスへの適切なアクセス権があることを確認してください。

## Flink で Iceberg を使用するための考慮事項

AWS Glue を Iceberg のカタログとして使用する場合は、テーブルを作成するデータベースが AWS Glue に存在することを確認してください。などのサービスを使用して AWS Lake Formation いて、カタログをロードできない場合は、コマンドを実行するためのサービスへの適切なアクセス権があることを確認してください。

## Hive で Iceberg を使用するための考慮事項

- Iceberg は以下のクエリタイプをサポートしています。
  - テーブルの作成
  - Drop table
  - Insert into table
  - Read table
- DML (データ操作言語MapReduce) オペレーションでは MR () 実行エンジンのみがサポートされており、Hive 3.1.3 では MR は非推奨です。
- AWS Glue Data Catalog は現在、Hive を使用した Iceberg ではサポートされていません。
- エラー処理の堅牢性が不十分です。設定に誤りがあっても、insert into クエリが正常に完了する可能性があります。ただし、メタデータの更新に失敗によりデータが失われる可能性があります。

## Iceberg リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Iceberg のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョン

については、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Iceberg バージョン情報

Amazon EMR リリースラベル	Iceberg バージョン	Iceberg でインストールされるコンポーネント
emr-7.1.0	1.4.3-amzn-0	Not available.
emr-7.0.0	1.4.2-amzn-0	Not available.
emr-6.15.0	1.4.0-amzn-0	Not available.
emr-6.14.0	1.3.1-amzn-0	Not available.
emr-6.13.0	1.3.0-amzn-1	Not available.
emr-6.12.0	1.3.0-amzn-0	Not available.
emr-6.11.1	1.2.0-amzn-0	Not available.
emr-6.11.0	1.2.0-amzn-0	Not available.
emr-6.10.1	1.1.0-amzn-0	Not available.
emr-6.10.0	1.1.0-amzn-0	Not available.
emr-6.9.1	0.14.1-amzn-0	Not available.
emr-6.9.0	0.14.1-amzn-0	Not available.
emr-6.8.1	0.14.0-amzn-0	Not available.
emr-6.8.0	0.14.0-amzn-0	Not available.
emr-6.7.0	0.13.1-amzn-0	Not available.
emr-6.6.0	0.13.1	Not available.
emr-6.5.0	0.12.0	Not available.

## バージョン別の Iceberg リリースノート

- [Amazon EMR 6.9.0 - Iceberg リリースノート](#)

### Amazon EMR 6.9.0 - Iceberg リリースノート

#### Amazon EMR 6.9.0 - Iceberg の変更点

型	説明
機能	Amazon EMR Flink と Iceberg の統合。
機能	Amazon EMR Hive と Iceberg の統合。
機能	Amazon FSx for Lustre で Iceberg メタデータ ファイルをキャッシュできるようになりました。これにより、クエリプランにかかる時間を短縮できます。
バックポート	<a href="#">PR 5050</a> : Flink 1.15: インライン挿入 SQL コメントの書き込みオプションがサポートされました。
バックポート	<a href="#">PR 5282</a> : AWS: 新しいデータファイルストリームを開いて PUT 再試行の失敗を修正しました。
バックポート	<a href="#">PR 5318</a> : Flink 1.15: FlinkSource と IcebergSource (FLIP-27) の間のギャップを埋め、Flink SQL で FLIP-27 ソースを使用するオプトイン設定を追加しました。
バックポート	<a href="#">PR 5344</a> : Flink 1.14: FlinkSource と IcebergSource (FLIP-27) の間のギャップを埋め、Flink SQL で FLIP-27 ソースを使用するオプトイン設定を追加しました。

型	説明
バックポート	<a href="#">PR 5393</a> : Flink 1.14、1.15: FLIP-27 ソースリーダーで Iceberg MetricContext を Flink メトリクスに変換することは避けてください。
バックポート	<a href="#">PR 5401</a> : Flink 1.14、1.15: FLIP-27 ソースリーダーメトリクスの PR #5393 の IcebergSourceReader グループを見逃しました。
バックポート	<a href="#">PR 5679</a> : Spark 3.2、3.3: MergeRows ノードの null 可能性の伝播を修正しました。
バックポート	<a href="#">PR 5860</a> : Spark 3.3: Date パーティションテーブル RewriteManifestProcedure で実行 QueryFailure した場合の修正。
バックポート	<a href="#">PR 5880</a> : Spark 3.3: merge-on-read 射影の null 性を修正しました。
バックポート	<a href="#">PR 5917</a> : Spark 3.2: merge-on-read 射影の null 性を修正しました。

# Amazon EMR の Jupyter Notebook

[Jupyter Notebook](#) はオープンソースのウェブアプリケーションで、ライブコード、方程式、視覚化、音声テキストを含むドキュメントの作成や共有に使用することができます。Amazon EMR には、Jupyter Notebook を使用する 3 つのオプションがあります。

## トピック

- [EMR Studio](#)
- [Jupyter ノートブックに基づく Amazon EMR Notebooks](#)
- [JupyterHub](#)

## EMR Studio

Amazon EMR Studio は、Amazon EMR クラスターで実行される、フルマネージド型の [Jupyter Notebook](#) 用のウェブベースの統合開発環境 (IDE) です。チームが R、Python、Scala、および PySpark で記述されたアプリケーションの開発、視覚化、およびデバッグを行うために EMR Studio をセットアップできます。

Amazon EMR で Jupyter Notebook を使用する場合は、EMR Studio を使用することをお勧めします。詳細については、「Amazon EMR 管理ガイド」の「[EMR Studio](#)」を参照してください。

## Jupyter ノートブックに基づく Amazon EMR Notebooks

EMR Notebooks は Amazon EMR コンソールに組み込まれた [Jupyter Notebook](#) 環境です。これを使用して、Jupyter Notebooks を素早く作成し、それらを Spark クラスターにアタッチして、コンソールで Jupyter Notebook エディタを開き、クエリやコードをリモートで実行できます。EMR Notebooks はクラスターとは別個に Amazon S3 に保存されるため、耐久性の高いストレージ、迅速なアクセス、柔軟性が確保されます。複数のノートブックを開き、1 つのクラスターに複数のノートブックをアタッチし、異なるクラスターでノートブックを再利用できます。

詳細については、「Amazon EMR 管理ガイド」の「[EMR notebooks](#)」を参照してください。

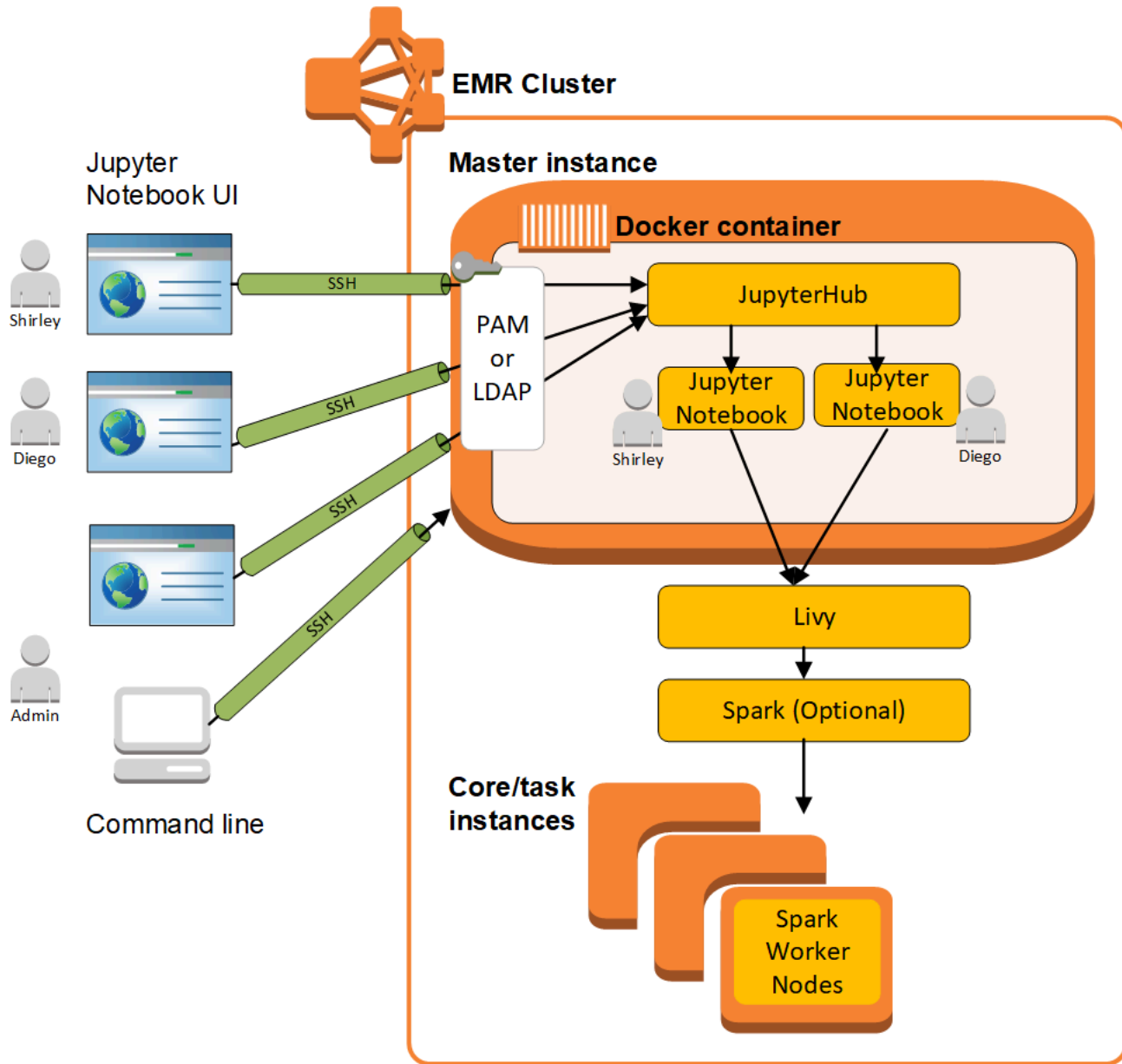
## JupyterHub

[Jupyter Notebook](#) はオープンソースのウェブアプリケーションで、ライブコード、方程式、視覚化、音声テキストを含むドキュメントの作成や共有に使用することができます。[JupyterHub](#) では、

単一ユーザー Jupyter Notebook サーバーの複数のインスタンスをホストできます。を使用してクラスターを作成すると JupyterHub、Amazon EMR はクラスターのマスターノードに Docker コンテナを作成します。JupyterHub、Jupyter に必要なすべてのコンポーネント、およびコンテナ内で実行される [Sparkmagic](#)。

Sparkmagic は Jupyter Notebook が Amazon EMR で実行されている [Apache Spark](#) (Spark の REST サーバーである [Apache Livy](#) を使用) と連携できるようにするカーネルのライブラリです。でクラスターを作成すると、Spark と Apache Livy が自動的にインストールされます JupyterHub。Jupyter 用のデフォルトの Python 3 カーネルは、PySpark 3 PySpark、および Sparkmagic で使用できる Spark カーネルとともに使用できます。こうしたカーネルを使用して臨時的 Spark コードや、Python、Scala を使ったインタラクティブな SQL クエリを実行できます。Docker コンテナ内で追加のカーネルを手動で追加することもできます。詳細については、「[追加のカーネルとライブラリをインストールする](#)」を参照してください。

次の図は、ノートブックユーザーと管理者に対応する認証方法を使用して、Amazon EMR JupyterHub 上の のコンポーネントを示しています。詳細については、「[Jupyter Notebook のユーザーと管理者を追加する](#)」を参照してください。



次の表に、Amazon EMR 7.x シリーズの最新リリース JupyterHub に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します JupyterHub。

このリリース JupyterHub で と共にインストールされるコンポーネントのバージョンについては、[「リリース 7.1.0 コンポーネントバージョン」](#)を参照してください。



## JupyterHub emr-7.1.0 のバージョン情報

Amazon EMR リリースラベル	JupyterHub バージョン	でインストールされるコンポーネント JupyterHub
emr-7.1.0	JupyterHub 1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

次の表に、Amazon EMR 6.x シリーズの最新リリース JupyterHub に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します JupyterHub。

このリリース JupyterHub で と共にインストールされるコンポーネントのバージョンについては、[「リリース 6.15.0 コンポーネントバージョン」](#)を参照してください。

## JupyterHub emr-6.15.0 のバージョン情報

Amazon EMR リリースラベル	JupyterHub バージョン	でインストールされるコンポーネント JupyterHub
emr-6.15.0	JupyterHub 1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR リリースラベル	JupyterHub バージョン	でインストールされるコンポーネント JupyterHub
		, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

次の表に、Amazon EMR 5.x シリーズの最新リリース JupyterHub に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します JupyterHub。

このリリース JupyterHub で と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

#### JupyterHub emr-5.36.2 のバージョン情報

Amazon EMR リリースラベル	JupyterHub バージョン	でインストールされるコンポーネント JupyterHub
emr-5.36.2	JupyterHub 1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR JupyterHub の に含まれている Python 3 カーネルは 3.6.4 です。

jupyterhub コンテナ内にインストールされているライブラリは Amazon EMR リリースバージョンと Amazon EC2 AMI バージョンで異なる場合があります。

## conda を使用してインストール済みのライブラリをリストにする

- マスターノードのコマンドラインで次のコマンドを実行します。

```
sudo docker exec jupyterhub bash -c "conda list"
```

## pip を使用してインストール済みのライブラリをリストにする

- マスターノードのコマンドラインで次のコマンドを実行します。

```
sudo docker exec jupyterhub bash -c "pip freeze"
```

## トピック

- [でクラスターを作成する JupyterHub](#)
- [Amazon EMR でを使用する場合 JupyterHub の考慮事項](#)
- [の設定 JupyterHub](#)
- [Amazon S3 でノートブックの永続性を設定するには](#)
- [マスターノードとノートブックサーバーに接続する](#)
- [JupyterHub 設定と管理](#)
- [Jupyter Notebook のユーザーと管理者を追加する](#)
- [追加のカーネルとライブラリをインストールする](#)
- [JupyterHub リリース履歴](#)

## でクラスターを作成する JupyterHub

、また、または Amazon EMR API JupyterHub を使用して AWS Management Console AWS Command Line Interface Amazon EMR クラスターを作成できます。ステップ完了後に自動終了するオプションでクラスターを作成していないことを確認します (AWS CLIの --auto-terminate オプション)。また、管理者とノートブックユーザーがクラスターの作成時に使用するキーペアにアクセスできることを確認します。詳細は、「Amazon EMR 管理ガイド」の「[SSH 認証情報にキーペアを使用する](#)」を参照してください。

## コンソール JupyterHub を使用して でクラスターを作成する

Amazon EMR コンソールのアドバンスドオプションを使用して JupyterHub がインストールされたクラスターを作成するには、次の手順に従います。

Amazon EMR コンソールを使用して JupyterHub がインストールされた Amazon EMR クラスターを作成するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Software Configuration (ソフトウェア設定)] で次を実行します。
  - リリースで、emr-5.36.2 を選択し、 を選択します JupyterHub。
  - Spark を使用する場合、Spark SQL のメタストアとして AWS Glue データカタログを使用するには、「Use for Spark table metadata 」を選択します。詳細については、「[AWS Glue データカタログを Spark SQL のメタストアとして使用する](#)」を参照してください。
  - [Edit software settings (ソフトウェア設定の編集)] には [Enter configuration (設定の入力)] を選択し、値を指定するか [Load JSON from S3 (S3 からの JSON のロード)] を選び JSON 設定ファイルを指定します。詳細については、「[の設定 JupyterHub](#)」を参照してください。
4. クラスターを作成した時点で [Add steps (optional) (ステップの追加 (オプション))] で実行するステップを設定する場合は、[Auto-terminate cluster after the last step is completed (最後のステップが完了したらクラスターを自動終了する)] を選択せずに、[Next (次へ)] を選びます。
5. [Hardware Configuration (ハードウェア構成)] オプション、[Next (次へ)] の順に選択します。詳細については、「Amazon EMR 管理ガイド」の「[クラスターハードウェアとネットワークを構成する](#)」を参照してください。
6. [General Cluster Settings (クラスターの全般設定)] オプションを選択し [Next (次へ)] を選びます。
7. [Security Options (セキュリティオプション)] でキーペアを指定し、[Create Cluster (クラスターの作成)] を選択します。

## JupyterHub を使用して クラスターを作成する AWS CLI

でクラスターを起動するには JupyterHub、`aws emr create-cluster` コマンドを使用し、`--applications` オプションに `Name=JupyterHub` を指定します。次の例では、JupyterHub2 つの EC2 インスタンス (1 つのマスタースタンスと 1 つのコアインスタンス) を使用して Amazon EMR でクラスターを起動します。デバッグも有効になっています。`--log-uri` が指定する Amazon S3 ロケーションにログが保存されています。指定されたキーペアはクラスター内の Amazon EC2 インスタンスへのアクセスを提供します。

### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \  
--applications Name=JupyterHub --log-uri s3://MyBucket/MyJupyterClusterLogs \  
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes  
KeyName=MyKeyPair
```

## Amazon EMR で を使用する場合 JupyterHub の考慮事項

Amazon EMR JupyterHub で を使用する場合は、次の点を考慮してください。

### Warning

ユーザーノートブックとファイルはマスタースタンスのファイルシステムに保存されています。これはクラスター終了で保持されないエフェメラルストレージです。クラスターが終了すると、バックアップされていないデータは失われます。cron ジョブまたは別のアプリケーションに適したものを使用して、定期的にバックアップを行うことを推奨します。また、コンテナ内で行った設定変更はコンテナの再起動時に保持されない場合があります。カスタマイズをより再現しやすくするため、スクリプトを使用するかコンテナ設定の自動化を推奨します。

- Amazon EMR セキュリティ設定を使用してセットアップした Kerberos 認証はサポートされていません。
- [OAuthenticator](#) はサポートされていません。

## の設定 JupyterHub

クラスターマスターノードに接続し、設定ファイルを編集することで、Amazon EMR および個々のユーザーノートブック JupyterHub の の設定をカスタマイズできます。値を変更したら jupyterhub コンテナを再起動します。

次のファイルのプロパティを変更して、 JupyterHub と個々の Jupyter Notebook を設定します。

- `jupyterhub_config.py` - デフォルトでは、このファイルはマスターノードの `/etc/jupyter/conf/` ディレクトリに保存されています。詳細については、 JupyterHub ドキュメントの「[設定の基本](#)」を参照してください。
- `jupyter_notebook_config.py` - このファイルは、デフォルトで `/etc/jupyter/` ディレクトリに保存され、デフォルトとして jupyterhub コンテナにコピーされます。詳細については、 Jupyter Notebook のドキュメントで「[Config file and command line options](#)」を参照してください。

`jupyter-sparkmagic-conf` 設定分類を使用して Sparkmagic をカスタマイズすることもできます。これにより Sparkmagic の `config.json` ファイルの値が更新されます。使用可能な設定の詳細については、 の [example\\_config.json GitHub](#)を参照してください。Amazon EMR のアプリケーションで設定分類を使用する方法については「[アプリケーションの設定](#)」を参照してください。

次の例では、Sparkmagic 設定分類設定 `MyJupyterConfig.json` の ファイルを参照して AWS CLI、 を使用してクラスターを起動します。

### Note

読みやすくするために、Linux 行連続文字 (`\`) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --use-default-roles --release-label emr-5.14.0 \  
--applications Name=Jupyter --instance-type m4.xlarge --instance-count 3 \  
--ec2-attributes KeyName=MyKey,SubnetId=subnet-1234a5b6 --configurations file://  
MyJupyterConfig.json
```

`MyJupyterConfig.json` のサンプルコンテンツは次のようになります。

```
[
  {
    "Classification": "jupyter-sparkmagic-conf",
    "Properties": {
      "kernel_python_credentials" : "{\"username\": \"diego\", \"base64_password\": \"mypass\", \"url\": \"http://localhost:8998\", \"auth\": \"None\"}"
    }
  }
]
```

### Note

Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## Amazon S3 でノートブックの永続性を設定するには

ユーザーによって保存されたノートブックが JupyterHub クラスター EC2 インスタンスのエフェメラルストレージの外部で Amazon S3 に保持されるように、Amazon EMR でクラスターを設定できます。

クラスターの作成時に、`jupyter-s3-conf` 設定分類を使用して Amazon S3 の永続性を指定します。詳細については、「[アプリケーションの設定](#)」を参照してください。

`s3.persistence.enabled` プロパティを使用して Amazon S3 の永続性を有効にすることに加えて、`s3.persistence.bucket` プロパティを使用してノートブックが保存される Amazon S3 でバケットを指定します。各ユーザーのノートブックは、指定したバケットの `jupyter/jupyterhub-user-name` フォルダに保存されます。バケットは Amazon S3 に存在する必要があります。また、クラスターの作成時に指定した EC2 インスタンスプロファイルのロールには、バケットへのアクセス権限が必要です (デフォルトでは、ロールは `EMR_EC2_DefaultRole` です)。詳細については、「[サービスへの Amazon EMR アクセス許可の IAM ロールを設定する AWS](#)」を参照してください。

同じ設定分類プロパティを使用して新しいクラスターを起動するとき、ユーザーは保存した場所からの内容を含むノートブックを開くことができます。

Amazon S3 を有効にした状態でノートブックでモジュールとしてファイルをインポートすると、ファイルは Amazon S3 にアップロードされます。Amazon S3 永続化を有効にせずにファイルをインポートすると、JupyterHub コンテナにアップロードされます。

次の例では、Amazon S3 の永続性を有効にします。ユーザーが保存したノートブックは、各ユーザーの `s3://MyJupyterBackups/jupyter/jupyterhub-user-name` フォルダに保存されます。`jupyterhub-user-name` は `diego` などのユーザー名です。

```
[
  {
    "Classification": "jupyter-s3-conf",
    "Properties": {
      "s3.persistence.enabled": "true",
      "s3.persistence.bucket": "MyJupyterBackups"
    }
  }
]
```

## マスターノードとノートブックサーバーに接続する

JupyterHub 管理者とノートブックユーザーは、SSH トンネルを使用してクラスターマスターノードに接続し、マスターノード JupyterHub が提供するウェブインターフェイスに接続する必要があります。SSH トンネルの設定やウェブ接続をプロキシするためにトンネルを使用する場合の詳細については、「Amazon EMR 管理ガイド」の「[クラスターに接続する](#)」を参照してください。

デフォルトでは、JupyterHub Amazon EMR はマスターノードのポート 9443 から使用できます。内部 JupyterHub プロキシは、ポート 9443 を介してノートブックインスタンスも提供 JupyterHub し、Jupyter ウェブインターフェイスには、次のパターンの URL を使用してアクセスできます。

`https://MasterNodeDNS:9443`

`c.JupyterHub.port` プロパティ (`jupyterhub_config.py` ファイル) を使用して別のポートを指定することができます。詳細については、JupyterHub ドキュメントの「[ネットワークの基本](#)」を参照してください。

デフォルトでは、JupyterHub Amazon EMR は HTTPS を使用した SSL 暗号化に自己署名証明書を使用します。接続が完了すると、自己署名証明書を信頼することを確認するメッセージが表示されます。ご自分の信頼済みの証明書とキーを使用できます。デフォルトの証明書ファイル、`server.crt`、キーファイル `server.key` をマスターノードの `/etc/jupyter/conf/`



ディレクトリでご自分の証明書とキーファイルに置き換えます。c.JupyterHub.ssl\_key と c.JupyterHub.ssl\_cert プロパティ (jupyterhub\_config.py ファイル) を使用して SSL マテリアルを指定します。詳細については、JupyterHub ドキュメントの「[セキュリティ設定](#)」を参照してください。jupyterhub\_config.py を更新したら、コンテナを再起動します。

## JupyterHub 設定と管理

JupyterHub および関連コンポーネントは、Ubuntu オペレーティングシステム jupyterhub を実行するという名前の Docker コンテナ内で実行されます。コンテナ内で実行するコンポーネントを管理する方法はいくつかあります。

### Warning

コンテナ内で行うカスタマイズはコンテナの再起動時に保持されない場合があります。カスタマイズをより再現しやすくするため、スクリプトを使用するかコンテナ設定の自動化を推奨します。

## コマンドラインを使用して管理する

SSH を使用してマスターノードに接続する場合は、Docker コマンドラインインターフェイス (CLI) を使用してコマンドを実行できます。その場合はコンテナを名前または ID で指定します (jupyterhub)。たとえば、`sudo docker exec jupyterhub command` はオペレーティングシステムまたはコンテナ内で実行されるアプリケーションが認識するコマンドを実行します。このメソッドを使用し、オペレーティングシステムにユーザーを追加して Docker コンテナ内に追加アプリケーションやライブラリをインストールすることができます。たとえば、デフォルトのコンテナイメージにはパッケージのインストールの Conda が含まれているので、マスターノードラインで次のコマンドを実行し、コンテナ内でアプリケーション、Keras をインストールすることができます。

```
sudo docker exec jupyterhub conda install keras
```

## ステップを送信して管理する

ステップはクラスターに作業を送信する方法です。クラスターを起動するときにステップを送信、または実行中のクラスターにステップを送信することができます。command-runner.jar を使用し、コマンドラインで実行するコマンドをステップとして送信できます。詳細については「[Amazon EMR 管理ガイド](#)」の「[CLI およびコンソールを使用した手順の作業](#)」と「[Amazon EMR クラスターでのコマンドとスクリプトの実行](#)」を参照してください。

例えば、ローカルコンピュータで次の AWS CLI コマンドを使用して、前の例のマスターノードのコマンドラインから行ったのと同じ方法で Keras をインストールできます。

```
aws emr add-steps --cluster-id MyClusterID --steps Name="Command Runner",Jar="command-runner.jar",Args="/usr/bin/sudo","/usr/bin/docker","exec","jupyterhub","conda","install","keras"
```

一連のステップをスクリプトにし、そのスクリプトを Amazon S3 にアップロードすると、クラスターの作成時またはステップとしてスクリプトを追加する場合に `script-runner.jar` を使用してスクリプトを実行できます。詳細については、「[Amazon EMR クラスターでのコマンドとスクリプトの実行](#)」を参照してください。例については、[the section called “例: Bash スクリプトで複数のユーザーを追加する”](#)を参照してください。

## REST API を使用して管理する

Jupyter、JupyterHub、およびの HTTP プロキシは、リクエストの送信に使用できる REST APIs JupyterHub を提供します。にリクエストを送信するには JupyterHub、リクエストで API トークンを渡す必要があります。マスターノードのコマンドラインから `curl` コマンドを使用して、REST コマンドを実行できます。詳細については、以下のリソースを参照してください。

- のドキュメントで [JupyterHubの REST API を使用する](#)。これには JupyterHubAPI トークンを生成する手順が含まれています。
- の [Jupyter Notebook サーバー API](#) GitHub
- [configurable-http-proxy](#) 上の GitHub

次の例は、の REST API を使用してユーザーのリスト JupyterHub を取得する方法を示しています。コマンドは、以前に生成された管理トークンを渡し、デフォルトのポート 9443 をに使用し JupyterHub、表示しやすくするために出力を `jq` にパイプします。

```
curl -XGET -s -k https://$HOST:9443/hub/api/users \  
-H "Authorization: token $admin_token" | jq .
```

## Jupyter Notebook のユーザーと管理者を追加する

ユーザーがノートブックを作成し、オプションでを管理 JupyterHub できるように、2 つの方法のいずれかを使用してを認証できます JupyterHub。最も簡単な方法は、JupyterHubのプラグブル認証モジュール (PAM) を使用することです。さらに、JupyterHub Amazon EMR では、Microsoft Active

Directory サーバーなどの LDAP サーバーからユーザー ID を取得するための [LDAP 認証プラグイン JupyterHub](#) をサポートしています。各認証方法でユーザーを追加する場合の手順や例については、こちらのセクションをご覧ください。

JupyterHub Amazon EMR のには、管理者権限を持つデフォルトユーザーがあります。ユーザー名は `jovyan` でパスワードは `jupyter` です。このユーザーを管理者の権限を持つ別のユーザーに置き換えることを強く推奨します。これを行うには、クラスターの作成時にステップを使用するか、クラスターの実行中にマスターノードに接続します。

## トピック

- [PAM 認証の使用](#)
- [LDAP 認証の使用](#)
- [ユーザー偽装](#)

## PAM 認証の使用

Amazon EMR JupyterHub での PAM ユーザーの作成は、2 ステップのプロセスです。最初のステップは、マスターノードで `jupyterhub` コンテナを実行しているオペレーティングシステムにユーザーを追加し、対応するユーザーのホームディレクトリに各ユーザーを追加することです。2 番目のステップは、これらのオペレーティングシステムユーザーをユーザーとして JupyterHub に追加することです。これは、`sudo` でホワイトリストに登録と呼ばれるプロセスです JupyterHub。JupyterHub ユーザーが追加されたら、JupyterHub URL に接続し、オペレーティングシステムの認証情報を提供してアクセスできます。

ユーザーがログインすると、はそのユーザーのノートブックサーバーインスタンス JupyterHub を開き、マスターノードのユーザーのホームディレクトリに保存されます `/var/lib/jupyter/home/username`。ノートブックサーバーインスタンスが存在しない場合、はユーザーのホームディレクトリにノートブックインスタンスを JupyterHub スポーンします。以下のセクションでは、オペレーティングシステムとにユーザーを個別に追加し JupyterHub、その後複数のユーザーを追加する基本的な `bash` スクリプトを追加する方法を示します。

### コンテナにオペレーティングシステムユーザーを追加する

次の例は、最初にコンテナ内の `useradd` コマンドを使用して 1 人のユーザー (`diego`) を追加してから、そのユーザー用のホームディレクトリを作成します。2 番目のコマンドは `chpasswd` を使用して、このユーザーに `diego` のパスワードを確立します。SSH で接続中にマスターノードコマンドラインでコマンドが実行されます。[ステップを送信して管理する](#) で説明されているように、ステップを使用してこうしたコマンドを実行することもできます。

```
sudo docker exec jupyterhub useradd -m -s /bin/bash -N diego
sudo docker exec jupyterhub bash -c "echo diego:diego | chpasswd"
```

## JupyterHub ユーザーの追加

の管理パネル JupyterHub または REST API を使用して、ユーザーと管理者、またはユーザーのみを追加できます。

の管理者パネルを使用してユーザーと管理者を追加するには JupyterHub

1. SSH を使用してマスターノードに接続し、管理者権限を持つ ID を使用して `https://MasterNodeDNS:9443` にログインします。
2. [Control Panel (コントロールパネル)]、[Admin (管理者)] の順に選択します。
3. [User (ユーザー)] を選択し、[Add Users (ユーザーの追加)] または [Admin (管理者)]、[Add Admins (管理者の追加)] を選択します。

REST API を使用してユーザーを追加する

1. SSH を使用してマスターノードに接続し、次のコマンドをマスターノードで使用するか、コマンドをステップとして実行します。
2. 管理トークンを取得して API リクエストを行い、次のステップ `AdminToken` で をそのトークンに置き換えます。
3. 次のコマンドを使用して、 をコンテナ内に作成されたオペレーティングシステムユーザー `UserName` に置き換えます。

```
curl -XPOST -H "Authorization: token AdminToken" "https://$(hostname):9443/hub/api/users/UserName"
```

### Note

ウェブインターフェイスに JupyterHub 初めてログインすると、管理者 JupyterHub 以外のユーザーとして自動的に追加されます。

## 例: Bash スクリプトで複数のユーザーを追加する

次のサンプル bash スクリプトは、このセクションの前のステップを結び付けて、複数の JupyterHub ユーザーを作成します。スクリプトは直接マスターノードで実行したり、Amazon S3 にアップロードしてステップとして実行することができます。

スクリプトは最初にユーザー名の配列を確立し、jupyterhub token コマンドを使用してデフォルトの管理者 (jovyan) に API トークンを作成します。そうすると、jupyterhub コンテナにオペレーティングシステムユーザーを作成し、ユーザー名に対応する初期パスワードを割り当てます。最後に、REST API オペレーションを呼び出して、各ユーザーを作成します JupyterHub。以前スクリプトで生成したトークンを渡し、REST レスポンスを jq にパイプして見やすくします。

```
# Bulk add users to container and JupyterHub with temp password of username
set -x
USERS=(shirley diego ana richard li john mary anaya)
TOKEN=$(sudo docker exec jupyterhub /opt/conda/bin/jupyterhub token jovyan | tail -1)
for i in "${USERS[@]}";
do
    sudo docker exec jupyterhub useradd -m -s /bin/bash -N $i
    sudo docker exec jupyterhub bash -c "echo $i:$i | chpasswd"
    curl -XPOST --silent -k https://$(hostname):9443/hub/api/users/$i \
-H "Authorization: token $TOKEN" | jq
done
```

s3://mybucket/createjupyterusers.sh のような Amazon S3 の場所にスクリプトを保存します。ステップとして実行するために script-runner.jar を使えるようになります。

例: クラスターの作成時にスクリプトを実行する (AWS CLI)

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://MyBucket/MyJupyterClusterLogs \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
KeyName=MyKeyPair \
```

```
--steps Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
mybucket/createjupyterusers.sh"]
```

## 既存のクラスターでスクリプトを実行する (AWS CLI)

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr add-steps --cluster-id j-XXXXXXXX --steps Type=CUSTOM_JAR,\
Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
mybucket/createjupyterusers.sh"]
```

## LDAP 認証の使用

Lightweight Directory Access Protocol (LDAP) は、Active Directory または OpenLDAP サーバーのような LDAP との互換性があるディレクトリサーバープロバイダーで保存されているユーザーやコンピュータなどのリソースに対応するオブジェクトのクエリや変更を行うアプリケーションプロトコルです。Amazon EMR で [この LDAP 認証プラグイン JupyterHub](#) を使用して、ユーザー認証に LDAP を使用できます。JupyterHub プラグインは LDAP ユーザーのログインセッションを処理し、Jupyter にユーザー情報を提供します。これにより、ユーザーは LDAP 互換サーバーに保存されている ID の認証情報を使用して JupyterHub および ノートブックに接続できます。

このセクションのステップでは、の LDAP Authenticator プラグインを使用して LDAP をセットアップおよび有効にするステップについて説明します JupyterHub。このステップはマスターノードのコマンドラインに接続している状態で実行します。詳細については、「[マスターノードとノートブックサーバーに接続する](#)」を参照してください。

1. ホスト IP アドレス、ポート、バインディング名など、LDAP サーバーに関する情報を使用して LDAP 設定ファイルを作成します。
2. を変更/etc/jupyter/conf/jupyterhub\_config.pyして、の LDAP Authenticator プラグインを有効にします JupyterHub。
3. jupyterhub コンテナ内で LDAP を設定するスクリプトを作成して実行します。

4. ユーザーの LDAP をクエリし、各ユーザーのコンテナ内にホームディレクトリを作成します。ノートブックをホストするには、ホームディレクトリ JupyterHub が必要です。
5. 再起動するスクリプトを実行する JupyterHub

#### Important

LDAPを設定する前に、LDAP サーバーとクラスターマスターノードが必要に応じて通信できることを確認するため、ネットワークインフラストラクチャをテストしてください。TLS は通常、単純な TCP 接続でポート 389 を使用します。LDAP 接続が SSL を使用する場合の一般的な SSL の TCP ポートは 636 です。

### LDAP 設定ファイルを作成する

以下の例では、次のプレースホルダー設定の値を使用します。これらの値を実装と一致するパラメータに置き換えます。

- LDAP サーバーはバージョン 3 を実行していて、ポート 389 で使用可能になっています。これは標準的な LDAP の非 SSL ポートです。
- ベース識別子名 (DN) は `dc=example, dc=org` です。

テキストエディタを使用して、次のようなコンテンツを含む [ldap.conf](#) ファイルを作成します。LDAP 実装に適切な値を使用します。`host` を IP アドレスまたは LDAP サーバーの解決可能なホスト名に置き換えます。

```
base dc=example,dc=org
uri ldap://host
ldap_version 3
binddn cn=admin,dc=example,dc=org
bindpw admin
```

### の LDAP 認証プラグインを有効にする JupyterHub

テキストエディタを使用して `/etc/jupyter/conf/jupyterhub_config.py` ファイルを変更し、次のような [ldapauthenticator](#) プロパティを追加します。`host` を IP アドレスまたは LDAP サーバーの解決可能なホスト名に置き換えます。この例では、ユーザーオブジェクトが `people` という組織単位 (ou) にあり、先に `ldap.conf` を使って確立した識別名のコンポーネントを使用しているものとします。



```
c.JupyterHub.authenticator_class = 'ldapauthenticator.LDAPAuthenticator'  
c.LDAPAuthenticator.use_ssl = False  
c.LDAPAuthenticator.server_address = 'host'  
c.LDAPAuthenticator.bind_dn_template = 'cn={username},ou=people,dc=example,dc=org'
```

## コンテナ内で LDAP を設定する

テキストエディタを使用して、次のコンテンツを含む bash スクリプトを作成します。

```
#!/bin/bash  
  
# Uncomment the following lines to install LDAP client libraries only if  
# using Amazon EMR release version 5.14.0. Later versions install libraries by default.  
# sudo docker exec jupyterhub bash -c "sudo apt-get update"  
# sudo docker exec jupyterhub bash -c "sudo apt-get -y install libnss-ldap libpam-ldap  
  ldap-utils nscd"  
  
# Copy ldap.conf  
sudo docker cp ldap.conf jupyterhub:/etc/ldap/  
sudo docker exec jupyterhub bash -c "cat /etc/ldap/ldap.conf"  
  
# configure nss switch  
sudo docker exec jupyterhub bash -c "sed -i 's/\(^passwd.*\)/\1 ldap/g' /etc/  
nsswitch.conf"  
sudo docker exec jupyterhub bash -c "sed -i 's/\(^group.*\)/\1 ldap/g' /etc/  
nsswitch.conf"  
sudo docker exec jupyterhub bash -c "sed -i 's/\(^shadow.*\)/\1 ldap/g' /etc/  
nsswitch.conf"  
sudo docker exec jupyterhub bash -c "cat /etc/nsswitch.conf"  
  
# configure PAM to create home directories  
sudo docker exec jupyterhub bash -c "echo 'session required                pam_mkhomedir.so  
  skel=/etc/skel umask=077' >> /etc/pam.d/common-session"  
sudo docker exec jupyterhub bash -c "cat /etc/pam.d/common-session"  
  
# restart nscd service  
sudo docker exec jupyterhub bash -c "sudo service nscd restart"  
  
# Test  
sudo docker exec jupyterhub bash -c "getent passwd"  
  
# Install ldap plugin
```



```
sudo docker exec jupyterhub bash -c "pip install jupyterhub-ldapauthenticator"
```

マスターノードにスクリプトを保存し、マスターノードのコマンドラインから実行します。たとえば、`configure_ldap_client.sh` として保存されているスクリプトでファイルを実行可能にします。

```
chmod +x configure_ldap_client.sh
```

スクリプトを実行します。

```
./configure_ldap_client.sh
```

### Active Directory に属性を追加する

各ユーザーを検索してデータベースに適切なエントリを作成するには、JupyterHub Docker コンテナで Active Directory の対応するユーザーオブジェクトに次の UNIX プロパティが必要です。詳細については、記事「[Clarification regarding the status of identity management for Unix \(IDMU\) and NIS server role in Windows Server 2016 technical preview and beyond](#)」(Windows Server 2016テクニカルプレビュー以降でのUnix ( IDMU ) およびNISサーバーの役割のID管理のステータスに関する説明)、のセクション「How do I continue to edit the GID/UID RFC 2307 attributes now that the Unix Attributes Plug-in is no longer available for the Active Directory Users and Computers MMC snap-in?」(Unix属性プラグインがActiveDirectoryユーザーとコンピューターMMCスナップインで使用できなくなった場合、GID / UID RFC 2307属性を編集し続けるにはどうすればよいですか?) を参照してください。

- `homeDirectory`

これは、ユーザーのホームディレクトリへの場所 (通常は `/home/username`) です。

- `gidNumber`

これは、別のユーザーがすでに使用していない 60000 より大きい値です。使用中の GID については、`etc/passwd` ファイルを確認してください。

- `uidNumber`

これは、別のグループがすでに使用していない 60000 より大きい値です。使用中の UID については、`etc/group` ファイルを確認してください。

- `uid`

これは `username` と同じです。

## ユーザーのホームディレクトリを作成する

JupyterHub では、LDAP ユーザーを認証し、インスタンスデータを保存するために、コンテナ内のホームディレクトリが必要です。次の例は LDAP ディレクトリの shirley と diego という 2 人のユーザーを示しています。

まず、次の例のように [ldapsearch](#) を使用して、各ユーザーの ID とグループ ID 情報に LDAP サーバーをクエリします。*host* は LDAP サーバーの IP アドレスまたは解決可能なホスト名に置き換わります。

```
ldapsearch -x -H ldap://host \  
-D "cn=admin,dc=example,dc=org" \  
-w admin \  
-b "ou=people,dc=example,dc=org" \  
-s sub \  
"(objectclass=*)" uidNumber gidNumber
```

ldapsearch コマンドが shirley と diego といった 2 人のユーザーに見られるような LDIF 形式のレスポンスを返します。

```
# extended LDIF  
  
# LDAPv3  
# base <ou=people,dc=example,dc=org> with scope subtree  
# filter: (objectclass=*)  
# requesting: uidNumber gidNumber sn  
  
# people, example.org  
dn: ou=people,dc=example,dc=org  
  
# diego, people, example.org  
dn: cn=diego,ou=people,dc=example,dc=org  
sn: B  
uidNumber: 1001  
gidNumber: 100  
  
# shirley, people, example.org  
dn: cn=shirley,ou=people,dc=example,dc=org  
sn: A  
uidNumber: 1002  
gidNumber: 100
```

```
# search result
search: 2
result: 0 Success

# numResponses: 4
# numEntries: 3
```

レスポンスからの情報を使用し、コンテナ内のコマンドを実行して各ユーザーの共通名 (cn) にホームディレクトリを作成します。uidNumber と gidNumber を使用して、そのユーザーのホームディレクトリの所有権を修正します。その操作を次の例のコマンドがユーザーの *shirley* に行います。

```
sudo docker container exec jupyterhub bash -c "mkdir /home/shirley"
sudo docker container exec jupyterhub bash -c "chown -R $uidNumber /home/shirley"
sudo docker container exec jupyterhub bash -c "sudo chgrp -R $gidNumber /home/shirley"
```

### Note

の LDAP 認証 JupyterHub 機能は、ローカルユーザーの作成をサポートしていません。詳細については、「[LDAP authenticator configuration note on local user creation](#)」を参照してください。

ローカルユーザーを手動で作成するには、次のコマンドを使用します。

```
sudo docker exec jupyterhub bash -c "echo 'shirley:x:$uidNumber:$gidNumber:~/home/shirley:/bin/bash' >> /etc/passwd"
```

## コンテナを再起動する JupyterHub

以下のコマンドを実行して、jupyterhub コンテナを再起動します。

```
sudo docker stop jupyterhub
sudo docker start jupyterhub
```

## ユーザー偽装

Jupyter Notebook 内で実行されている Spark ジョブは、Amazon EMR での実行中に、複数のアプリケーションを走査します。例えば、ユーザーが Jupyter 内で実行する PySpark3 つのコードは Sparkmagic によって受信されます。Sparkmagic は HTTP POST リクエストを使用して Livy に送信し、YARN を使用してクラスターで実行する Spark ジョブを作成します。

デフォルトでは、このように送信された YARN ジョブは、ジョブを開始したユーザーとは関係なく、ユーザー livy として実行されます。ユーザー偽装を設定することで、ノートブックユーザーのユーザー ID を、YARN ジョブに関連付けられているユーザーとすることができます。各ユーザーによって開始されるジョブは、ユーザー shirley と関連付けられた diego と livy の両方によって開始するのではなく、それぞれ shirley および diego に関連付けられます。これにより、Jupyter の使用を監査し、組織内でアプリケーションを管理できます。

この設定がサポートされるのは、Sparkmagic から Livy への呼び出しが認証されていない場合のみです。Hadoop アプリケーションと Livy (Apache Knox Gateway など) 間で認証またはプロキシレイヤーを提供するアプリケーションはサポートされません。このセクションのユーザー偽装を設定する手順は、JupyterHub と Livy が同じマスターノードで実行されていることを前提としています。アプリケーションに別々のクラスターがある場合、[ステップ 3: ユーザーの HDFS ホームディレクトリを作成する](#) を変更し、HDFS ディレクトリが Livy マスターノードで作成されるようにする必要があります。

### ユーザー偽装を設定するステップ

- [ステップ 1: Livy を設定する](#)
- [ステップ 2: ユーザーを追加する](#)
- [ステップ 3: ユーザーの HDFS ホームディレクトリを作成する](#)

### ステップ 1: Livy を設定する

次の例に示すように、クラスターを作成して Livy のユーザー偽装を有効にする場合は、livy-conf および core-site 設定分類を使用します。設定分類を JSON として保存し、クラスターの作成時にその分類を参照するか、インラインで指定します。詳細については、「[アプリケーションの設定](#)」を参照してください。

```
[
  {
    "Classification": "livy-conf",
    "Properties": {
      "livy.impersonation.enabled": "true"
    }
  },
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.proxyuser.livy.groups": "*",
      "hadoop.proxyuser.livy.hosts": "*"
    }
  }
]
```

```
    }  
  }  
]
```

## ステップ 2: ユーザーを追加する

PAM または LDAP を使用して JupyterHub ユーザーを追加します。詳細については、「[PAM 認証の使用](#)」および「[LDAP 認証の使用](#)」を参照してください。

## ステップ 3: ユーザーの HDFS ホームディレクトリを作成する

ユーザーを作成するため、マスターノードに接続しました。マスターノードに接続中に、次のコンテンツをコピーして、スクリプトファイルに保存します。このスクリプトは、マスターノード上の JupyterHub ユーザーごとに HDFS ホームディレクトリを作成します。このスクリプトでは、デフォルトの管理者ユーザー ID *jovyan* を使用することが前提となっています。

```
#!/bin/bash  
  
CURL="curl --silent -k"  
HOST=$(curl -s http://169.254.169.254/latest/meta-data/local-hostname)  
  
admin_token() {  
  local user=jovyan  
  local pwd=jupyter  
  local token=$(($CURL https://$HOST:9443/hub/api/authorizations/token \  
    -d "{\"username\":\"$user\", \"password\":\"$pwd\"}" | jq ".token")  
  if [[ $token != null ]]; then  
    token=$(echo $token | sed 's//g')  
  else  
    echo "Unable to get Jupyter API Token."  
    exit 1  
  fi  
  echo $token  
}  
  
# Get Jupyter Admin token  
token=$(admin_token)  
  
# Get list of Jupyter users  
users=$(curl -XGET -s -k https://$HOST:9443/hub/api/users \  
  -H "Authorization: token $token" | jq '.[].name' | sed 's//g')  
  
# Create HDFS home dir
```

```
for user in ${users[@]};
do
  echo "Create hdfs home dir for $user"
  hadoop fs -mkdir /user/$user
  hadoop fs -chmod 777 /user/$user
done
```

## 追加のカーネルとライブラリをインストールする

Amazon EMR JupyterHub を使用してクラスターを作成すると、Jupyter 用のデフォルトの Python 3 カーネル PySpark と Sparkmagic 用の Spark カーネルが Docker コンテナにインストールされます。追加のカーネルをインストールできます。追加のライブラリやパッケージをインストールしてから、適切なシェルにインポートすることもできます。

### カーネルをインストールする

カーネルは Docker コンテナ内にインストールされます。これを簡単に行うには、インストールコマンドを使用して bash スクリプトを作成し、マスターノードに保存してから `sudo docker exec jupyterhub script_name` コマンドで jupyterhub コンテナ内にあるスクリプトを実行します。次の例のスクリプトはカーネルをインストールし、マスターノードでカーネルのライブラリをいくつかインストールすることで、後から Jupyter を使用してライブラリをインポートできるようにします。

```
#!/bin/bash

# Install Python 2 kernel
conda create -n py27 python=2.7 anaconda
source /opt/conda/envs/py27/bin/activate
apt-get update
apt-get install -y gcc
/opt/conda/envs/py27/bin/python -m pip install --upgrade ipykernel
/opt/conda/envs/py27/bin/python -m ipykernel install

# Install libraries for Python 2
/opt/conda/envs/py27/bin/pip install paramiko nltk scipy numpy scikit-learn pandas
```

コンテナ内でカーネルとライブラリをインストールするにはマスターノードへの接続を開き、スクリプトを `/etc/jupyter/install_kernels.sh` に保存してマスターノードのコマンドラインで次のコマンドを実行します。

```
sudo docker exec jupyterhub bash /etc/jupyter/install_kernels.sh
```

## ライブラリの使用と追加ライブラリのインストール

Python 3 用の機械学習ライブラリとデータサイエンスライブラリのコアセットは、Amazon EMR JupyterHub に がプリインストールされています。sudo docker exec jupyterhub bash -c "conda list" と sudo docker exec jupyterhub bash -c "pip freeze" を使用できません。

Spark ジョブがワーカーノードでライブラリを必要とする場合は、ブートストラップアクションを使用してクラスターの作成時にライブラリをインストールし、スクリプトを実行することを推奨します。ブートストラップアクションは、クラスターの作成時にすべてのクラスターノードで実行され、インストールプロセスが簡素化されます。クラスターを実行してから、コアまたはワーカーノードにライブラリをインストールするとオペレーションが複雑になります。このセクションの Python プログラムの例では、こうしたライブラリのインストール方法を示しています。

このセクションの例のブートストラップアクションと Python プログラムは、Amazon S3 に保存している bash スクリプトを使用してすべてのノードにライブラリをインストールします。

次の例で参照されているスクリプトは、pip を使用して paramiko、nltk、scipy、scikit-learn、pandas (Python 3 カーネル用) をインストールします。

```
#!/bin/bash

sudo python3 -m pip install boto3 paramiko nltk scipy scikit-learn pandas
```

スクリプトを作成したら、Amazon S3 の場所にアップロードします (例: s3://mybucket/install-my-jupyter-libraries.sh)。ブートストラップアクションまたは Python プログラムで使用できるようにするには、Amazon Simple Storage Service User Guide の [オブジェクトの Uploading](#) で詳細を確認してください。

を使用してクラスターを作成するときすべてのノードにライブラリをインストールするブートストラップアクションを指定するには AWS CLI

1. 前の例のようなスクリプトを作成し、Amazon S3 の場所に保存します。s3://mybucket/install-my-jupyter-libraries.sh の例を使用します。
2. を使用してクラスター JupyterHub を作成し、次の例に示すように、--bootstrap-actions オプションの Path 引数を使用してスクリプトの場所を指定します。

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://MyBucket/MyJupyterClusterLogs \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
KeyName=MyKeyPair \
--bootstrap-actions Path=s3://mybucket/install-my-jupyter-
libraries.sh,Name=InstallJupyterLibs
```

コンソールを使用してクラスターを作成する場合にすべてのノードでライブラリをインストールするためのブートストラップアクションを指定するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. アプリケーションの必要に応じて [Software and Steps (ソフトウェアとステップ)] および [Hardware (ハードウェア)] の設定を指定します。
4. [General Cluster Settings ( (クラスターの全般設定))] 画面で、[Bootstrap Actions (ブートストラップアクション)] を選択します。
5. [Add bootstrap action (ブートストラップアクションの追加)] で [Custom action (カスタムアクション)]、[Configure and add (設定と追加)] の順に選択します。
6. [Name] (名前) で、わかりやすい名前を入力します。スクリプトの場所には、スクリプトの Amazon S3 の場所を入力します (使用する例は s3://mybucket/.sh です ) install-my-jupyter-libraries。[Optional arguments (オプションの引数)] は空白のままにして、[Add (追加)] を選択します。
7. クラスターのその他の設定を指定し、[Next (次へ)] を選択します。
8. セキュリティ設定を指定し、[Create cluster (クラスターの作成)] を選択します。



## Example 実行中のクラスターのコアノードにライブラリをインストールする

Jupyter 内からマスターノードでライブラリのインストールを完了すると、実行中のコアノードでライブラリをさまざまな方法でインストールできます。次の例は、ローカルマシンで実行するように書き込まれた Python プログラムを示しています。Python プログラムをローカル AWS Systems Manager で実行すると、AWS-RunShellScript のを使用して、クラスターのコアノードにライブラリをインストールするこのセクションの前半で示したサンプルスクリプトを実行します。

```
import argparse
import time
import boto3

def install_libraries_on_core_nodes(cluster_id, script_path, emr_client, ssm_client):
    """
    Copies and runs a shell script on the core nodes in the cluster.

    :param cluster_id: The ID of the cluster.
    :param script_path: The path to the script, typically an Amazon S3 object URL.
    :param emr_client: The Boto3 Amazon EMR client.
    :param ssm_client: The Boto3 AWS Systems Manager client.
    """
    core_nodes = emr_client.list_instances(
        ClusterId=cluster_id, InstanceGroupTypes=["CORE"]
    )["Instances"]
    core_instance_ids = [node["Ec2InstanceId"] for node in core_nodes]
    print(f"Found core instances: {core_instance_ids}.")

    commands = [
        # Copy the shell script from Amazon S3 to each node instance.
        f"aws s3 cp {script_path} /home/hadoop",
        # Run the shell script to install libraries on each node instance.
        "bash /home/hadoop/install_libraries.sh",
    ]
    for command in commands:
        print(f"Sending '{command}' to core instances...")
        command_id = ssm_client.send_command(
            InstanceIds=core_instance_ids,
            DocumentName="AWS-RunShellScript",
            Parameters={"commands": [command]},
            TimeoutSeconds=3600,
        )["Command"]["CommandId"]
        while True:
```

```
# Verify the previous step succeeded before running the next step.
cmd_result = ssm_client.list_commands(CommandId=command_id)["Commands"][0]
if cmd_result["StatusDetails"] == "Success":
    print(f"Command succeeded.")
    break
elif cmd_result["StatusDetails"] in ["Pending", "InProgress"]:
    print(f"Command status is {cmd_result['StatusDetails']}, waiting...")
    time.sleep(10)
else:
    print(f"Command status is {cmd_result['StatusDetails']}, quitting.")
    raise RuntimeError(
        f"Command {command} failed to run. "
        f"Details: {cmd_result['StatusDetails']}"
    )

def main():
    parser = argparse.ArgumentParser()
    parser.add_argument("cluster_id", help="The ID of the cluster.")
    parser.add_argument("script_path", help="The path to the script in Amazon S3.")
    args = parser.parse_args()

    emr_client = boto3.client("emr")
    ssm_client = boto3.client("ssm")

    install_libraries_on_core_nodes(
        args.cluster_id, args.script_path, emr_client, ssm_client
    )

if __name__ == "__main__":
    main()
```

## JupyterHub リリース履歴

次の表に、Amazon EMR の各リリースバージョン JupyterHub に含まれる のバージョンと、アプリケーションと共にインストールされるコンポーネントを示します。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

## JupyterHub バージョン情報

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.36.2	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.1.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.0.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
		yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.15.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.14.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.13.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.12.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.11.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.11.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.10.1	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.10.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.9.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.9.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub



Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.8.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.8.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.7.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.36.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.36.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.6.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.35.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.5.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.4.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.3.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.3.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.2.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.2.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.1.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.1.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.0.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub



Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-6.0.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.34.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.33.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.33.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.32.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.32.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.31.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.31.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.30.2	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.30.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.30.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.29.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.28.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.28.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.27.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.27.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub



Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.26.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.25.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.24.1	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.24.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.23.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.23.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.22.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.2	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.21.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.20.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.20.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.19.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.19.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.18.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.18.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub



Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.17.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.17.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.17.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.16.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.16.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.15.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.15.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR リリースラベル	JupyterHub バージョン	と共にインストールされるコンポーネント JupyterHub
emr-5.14.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

# Apache Livy

Livy は、Spark を実行する EMR クラスターを、REST インターフェイスによって操作できます。REST インターフェイスまたは RPC クライアントライブラリを使って、Spark ジョブや Spark コードのスニペットを送信したり、同期的または非同期的に結果を取得したり、Spark コンテキストを管理したりすることができます。詳細については、「[Apache Livy ウェブサイト](#)」を参照してください。Livy は、Amazon EMR リリースバージョン 5.9.0 以降に含まれています。

Livy ウェブインターフェイスにアクセスするには、マスターノードへの SSH トンネルとプロキシ接続を設定します。詳細については、「[EMR クラスターでホストされているウェブサイトの表示](#)」を参照してください。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Livy のバージョンと、Amazon EMR で Livy と共にインストールされるコンポーネントを示しています。

このリリースで Livy と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Livy バージョン情報

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-7.1.0	Livy 0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Livy のバージョンと、Amazon EMR で Livy と共にインストールされるコンポーネントを示しています。

このリリースで Livy と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Livy バージョン情報

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.15.0	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Livy のバージョンと、Amazon EMR で Livy と共にインストールされるコンポーネントを示しています。

このリリースで Livy と共にインストールされるコンポーネントのバージョンについては、「[リリース 5.36.2 コンポーネントバージョン](#)」を参照してください。

#### emr-5.36.2 の Livy バージョン情報

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.36.2	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
		yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

## トピック

- [Apache Livy で HTTPS を有効にする](#)
- [Livy リリース履歴](#)

## Apache Livy で HTTPS を有効にする

1. 転送時の暗号化を有効にして Amazon EMR クラスターをプロビジョニングします。暗号化の詳細については、「[保管中と転送中のデータの暗号化](#)」を参照してください。
2. `livy_ssl.sh` というファイルを次の内容で作成します。

```
#!/bin/bash

KEYSTORE_FILE=`awk '/ssl.server.keystore.location/{getline; print}' /etc/hadoop/conf/ssl-server.xml | sed -e 's/<[^>]*>/g' | tr -d ' \t\n\r\f'`
KEYSTORE_PASS=`awk '/ssl.server.keystore.password/{getline; print}' /etc/hadoop/conf/ssl-server.xml | sed -e 's/<[^>]*>/g' | tr -d ' \t\n\r\f'`
KEY_PASS=`awk '/ssl.server.keystore.keypassword/{getline; print}' /etc/hadoop/conf/ssl-server.xml | sed -e 's/<[^>]*>/g' | tr -d ' \t\n\r\f'`

echo "livy.keystore $KEYSTORE_FILE
livy.keystore.password $KEYSTORE_PASS
livy.key-password $KEY_PASS" | sudo tee -a /etc/livy/conf/livy.conf >/dev/null

sudo systemctl restart livy-server.service
```



3. 次のスクリプトを Amazon EMR ステップとして実行します。このスクリプトは SSL を有効化するように `/etc/livy/conf/livy.conf` を変更します。

```
--steps '[{"Args":["s3://DOC-EXAMPLE-BUCKET/
livy_ssl.sh"],"Type":"CUSTOM_JAR","ActionOnFailure":"CONTINUE","Jar":"s3://
us-east-1.elasticmapreduce/libs/script-runner/script-
runner.jar","Properties":"","Name":"Custom JAR"}]'
```

4. Apache Livy サービスを再起動します。これにより変更が有効になります。Apache Livy を再起動する場合は、「[プロセスの停止と再起動](#)」を参照してください。
5. クライアントが HTTPS を使用して通信できるようになったことをテストします。例えば、ジョブを送信するには、次のコードを実行します。

```
curl -k -X POST --data '{"file": "local:///usr/lib/spark/examples/jars/spark-
examples.jar",
"className": "org.apache.spark.examples.SparkPi"}' \
-H "Content-Type: application/json" \
https://EMR_Master_Node_Host:8998/batches
```

HTTPS を正常に有効にすると、Livy はコマンドが受け付けられ、バッチジョブが送信されたことを示す応答を送信します。

```
{"id":1,"name":null,"owner":null,"proxyUser":null,"state":"starting","appId":null,"appInfo":
{"driverLogUrl":null,"sparkUiUrl":null},"log":["stdout: ", "\nstderr: ", "\nYARN
Diagnostics: "]}
```

## Livy リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Livy のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

## Livy バージョン情報

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.36.2	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.1.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.0.0	0.7.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
		yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.15.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.14.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.13.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.12.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.11.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.11.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.10.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.10.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.9.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.9.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.8.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.8.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx



Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.7.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.36.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.6.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.35.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.5.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.4.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.3.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.3.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.2.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.2.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.1.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.1.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.0.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-6.0.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.34.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx



Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.33.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.33.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.32.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.32.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.31.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.31.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.30.2	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.30.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.30.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.29.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.28.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.28.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.27.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.27.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.26.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.25.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx



Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.24.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.24.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.23.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.23.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.22.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.21.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.20.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.20.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.19.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.19.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.18.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.18.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.17.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.17.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server



Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.17.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.16.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.16.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.15.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.15.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.14.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.13.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.13.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.12.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.12.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.11.4	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server



Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.11.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.11.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.10.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.10.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.9.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR リリースラベル	Livy バージョン	Livy でインストールされるコンポーネント
emr-5.9.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

# Apache MXNet

Apache MXNet は、ニューラルネットワークやその他のディープラーニングアプリケーションを構築するために設計されたアクセラレーションライブラリです。MXNet は、一般的なワークフローを自動化し、数値計算を最適化します。MXNet により、線形代数オペレーションなどの低レベルの計算を実装することなく、ニューラルネットワークを設計することができます。MXNet は、Amazon EMR リリースバージョン 5.10.0 以降に含まれています。

詳細については、「[Apache MXNet ウェブサイト](#)」を参照してください。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている MXNet のバージョンと、Amazon EMR で MXNet と共にインストールされるコンポーネントを示しています。

このリリースで MXNet と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の MXNet バージョン情報

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-7.1.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている MXNet のバージョンと、Amazon EMR で MXNet と共にインストールされるコンポーネントを示しています。

このリリースで MXNet と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の MXNet バージョン情報

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.15.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている MXNet のバージョンと、Amazon EMR で MXNet と共にインストールされるコンポーネントを示しています。

このリリースで MXNet と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

## emr-5.36.2 の MXNet バージョン情報

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.36.2	MXNet 1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv

## MXNet リリース履歴

次の表は、Amazon EMR の各リリースに含まれている MXNet のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### MXNet バージョン情報

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.36.2	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.1.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.0.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
		server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.15.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.14.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.13.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv



Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.12.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.11.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.11.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.10.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.10.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.9.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.9.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.8.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.8.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.7.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.1	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.6.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.35.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.5.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.4.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.3.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.3.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.2.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.2.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.1.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-6.1.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.0.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.0.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv



Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.34.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.33.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.33.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.32.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.32.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.31.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.31.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.2	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.30.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.29.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.28.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.28.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.27.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.27.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.26.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.25.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.24.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.24.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.23.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.23.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.22.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.2	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv



Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.21.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.20.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.20.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.19.1	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.19.0	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.18.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.18.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.2	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.17.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.16.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.16.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.15.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.15.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.2	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.14.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.13.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.13.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.3	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.2	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.12.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.4	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet



Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.11.3	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.2	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Amazon EMR リリースラベル	MXNet バージョン	MXNet でインストールされるコンポーネント
emr-5.11.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.10.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.10.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

# Apache Oozie

Apache Oozie Workflow Scheduler を使用して、Hadoop ジョブを管理し、調整します。詳細については <http://oozie.apache.org/> をご覧ください。

Oozie ネイティブウェブインターフェイスは Amazon EMR でサポートされていません。Oozie のフロントエンドインターフェイスを使用するには、Hue Oozie アプリケーションを試します。詳細については、「[Hue](#)」を参照してください。Oozie は、Amazon EMR リリースバージョン 5.0.0 以降に含まれています。Oozie は以前のリリースにサンドボックスアプリケーションとして含まれています。詳細については、「[Amazon EMR 4.x リリースバージョン](#)」を参照してください。

作成日が 2018-08-11 の Amazon Linux AMI に基づくカスタム Amazon Linux AMI を使用すると、Oozie サーバーの起動に失敗します。Oozie を使用する場合は、作成日が異なる Amazon Linux AMI ID に基づいてカスタム AMI を作成します。次の AWS CLI コマンドを使用して、2018.03 バージョンのすべての HVM Amazon Linux AMIs のイメージ IDs とリリース日のリストを返すことができます。これにより、ベースとして適切な Amazon Linux AMI を選択できます。を us-west-2 などのリージョン識別子 MyRegion に置き換えます。

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`][?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Oozie のバージョンと、Amazon EMR で Oozie と共にインストールされるコンポーネントを示しています。

このリリースで Oozie と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Oozie バージョン情報

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-7.1.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Oozie のバージョンと、Amazon EMR で Oozie と共にインストールされるコンポーネントを示しています。

このリリースで Oozie と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Oozie バージョン情報

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.15.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Oozie のバージョンと、Amazon EMR で Oozie と共にインストールされるコンポーネントを示しています。

このリリースで Oozie と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

#### emr-5.36.2 の Oozie バージョン情報

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.36.2	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

#### トピック

- [Amazon RDS でリモートデータベースと Oozie を使用する](#)
- [Oozie 用の Java バージョンの設定](#)
- [Oozie リリース履歴](#)

## Amazon RDS でリモートデータベースと Oozie を使用する

デフォルトでは、Oozie のユーザー情報とクエリ履歴は、マスターノード上のローカルの MySQL データベースに保存されます。または、Amazon S3 に保存されている設定と、Amazon Relational Database Service(Amazon RDS) の MySQL データベースを使用して、Oozie が有効な 1 つ以上のクラスターを作成することができます。このようにすることで、Amazon EMR クラスターを実行し続

けなくとも、Oozie によって作成されたユーザー情報とクエリ履歴を永続化できます。設定ファイルは、Amazon S3 サーバーサイド暗号化機能を使って保存することをお勧めします。

最初に、Oozie のリモートデータベースを作成します。

外部の MySQL データベースを作成するには

1. Amazon RDS コンソール (<https://console.aws.amazon.com/rds/>) を開きます。
2. [DB インスタンスの起動] を選択します。
3. [MySQL]、[Select (選択)] の順に選択します。
4. [Multi-AZ Deployment and Provisioned IOPS Storage (マルチ AZ 配置とプロビジョンド IOPS ストレージ)] はデフォルトのままにし、[Next (次へ)] を選択します。
5. [インスタンスの仕様] はデフォルトのままにし、[Settings (設定)] を指定して、[Next (次へ)] を選択します。
6. [Configure Advanced Settings] ([詳細設定] の設定) ページで、適切なセキュリティグループとデータベース名を選択します。使用するセキュリティグループは、少なくともクラスターのマスターノードからのポート 3306 に対するインバウンド TCP アクセスを許可する必要があります。この時点でまだクラスターを作成していない場合、ポート 3306 への接続をすべてのホストに許可し、クラスターを起動した後にセキュリティグループを調整することができます。[DB インスタンスの起動] を選択します。
7. RDS ダッシュボードで、[Instances (インスタンス)] を選択し、先ほど作成したインスタンスを選択します。データベースが使用できる時、dbname、username、password、および RDS instance hostname をメモします。クラスターの作成および設定時に、この情報を使用します。

を使用してクラスターを起動するときに Oozie の外部 MySQL データベースを指定するには AWS CLI

AWS CLIを使用してクラスターを起動するときに Oozie の外部 MySQL データベースを指定するには、RDS インスタンスを作成する際にメモしておいた情報を、設定オブジェクトの `oozie-site` の設定に使用します。

#### Note

外部にある同一のデータベースを使用するが、各クラスターはクエリ履歴とユーザー情報を共有する複数のクラスターを作成できます。

- を使用して AWS CLI、Oozie がインストールされたクラスターを作成し、作成した外部データベースを使用し、データベースプロパティを指定する Oozie の設定分類を持つ設定ファイルを参照します。次の例では、Amazon S3 の設定ファイルとデータベース設定を指定する `myConfig.json` を参照して、Oozie がインストールされた状態でクラスターを作成します。

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --release-label emr-7.1.0 --applications Name=Oozie
Name=Spark Name=Hive \
--instance-type m5.xlarge --instance-count 3 \
--configurations https://s3.amazonaws.com/mybucket/myfolder/myConfig.json --use-
default-roles
```

`myConfig.json` ファイルのコンテンツの例を以下に示します。 **JDBC URL**、**username**、**password** は、RDS インスタンスの JDBC URL、ユーザー名、パスワードに置き換えます。

#### Important

JDBC URL には、データベース名をサフィックスとして含める必要があります。例えば、`jdbc:mysql://oozie-external-db.xxxxxxxxxx.us-east-1.rds.amazonaws.com:3306/dbname` です。

```
[{
  "Classification": "oozie-site",
  "Properties": {
    "oozie.service.JPIService.jdbc.driver": "org.mariadb.jdbc.Driver",
    "oozie.service.JPIService.jdbc.url": "JDBC URL",

    "oozie.service.JPIService.jdbc.username": "username",
    "oozie.service.JPIService.jdbc.password": "password"
  },
  "Configurations": []
}]
```

}]

## Oozie 用の Java バージョンの設定

Oozie は複数の Java Virtual Machine (JVM) プロセスを実行します。このページでは、各プロセスの Java バージョンを設定する方法について説明します。

- Oozie サーバー: oozie-env 分類の JAVA\_HOME を設定して、EmbeddedOozieServer の Java バージョンを更新します。
- Oozie Launcher AM: Oozie Launcher AM は、Hadoop、Hive などの適切なアプリケーションクライアントライブラリを呼び出す単一マッパ MR ジョブです。特に設定されていない限り、Oozie Launcher AM のランタイムバージョンは EMR クラスター内の Hadoop の Java ランタイムと同じです。Oozie Launcher AM 用の Java ランタイムを設定するには、このジョブの workflow.xml に次のプロパティを設定します。

```
<property>
  <name>mapred.child.env</name>
  <value>JAVA_HOME=/path/to/JAVA_HOME</value>
</property>
```

このプロパティにより、Oozie ジョブの Oozie Launcher AM が Hadoop に設定されている Java バージョンではなく、指定した Java バージョンで実行されるようになります。

- アプリケーションクライアント実行可能ファイル: Oozie Launcher AM はデフォルトでアプリケーションクライアントを起動するため、クライアント実行可能ファイルの Java ランタイムは Oozie Launcher AM と同じです。
- Oozie ジョブによって起動されるアプリケーション: 特に指定がない限り、Oozie ジョブによって起動される実際のアプリケーション JVM のランタイムバージョンは、EMR クラスター内の Hadoop 用 Java ランタイムと同じです。Oozie ジョブ (Spark または Hive アクション) でアプリケーションの起動に使用される Oozie ワークフローアクションのタイプに応じて、Oozie ジョブの workflow.xml で実際のアプリケーション JVM のデフォルト Java ランタイムを更新できます。

## Oozie リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Oozie のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョン



については、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

## Oozie バージョン情報

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.36.2	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-7.1.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.0.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされる コンポーネント
		cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.15.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.14.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.13.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.12.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.11.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.11.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.10.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.10.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.9.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.8.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn



Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.36.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.36.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.35.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.3.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.2.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.2.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.1.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.1.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-6.0.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.0.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.34.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.33.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn



Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.33.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.32.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.32.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.31.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.31.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.30.2	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.30.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.30.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.29.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.28.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.28.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.27.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.27.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.26.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.25.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.24.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn



Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.24.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.23.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.23.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.22.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.21.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.21.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.21.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.20.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.20.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.19.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.19.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.18.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.18.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.17.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.17.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.17.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn



Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.16.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.16.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.15.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.15.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.14.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.14.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.14.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.13.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.13.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.12.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.12.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.11.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn



Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.11.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.10.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.10.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.9.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.9.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.8.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.8.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.7.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.7.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.6.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.6.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.5.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn



Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.5.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.5.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.4.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.4.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.3.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.3.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.3.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.2.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.2.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.1.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.1.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.0.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn



Amazon EMR リリースラベル	Oozie バージョン	Oozie でインストールされるコンポーネント
emr-5.0.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

## バージョン別の Oozie リリースノート

### トピック

- [Amazon EMR 7.1.0 - Oozie の変更点](#)

## Amazon EMR 7.1.0 - Oozie の変更点

型	説明
Upgrade	Netty を 4.1.100.Final にアップグレードします。
Upgrade	Jetty を 9.4.53.v20231009 にアップグレードします。

# Apache Phoenix

Apache Phoenix は、OLTP と業務分析に使用され、Apache HBase バックリングストアと連携して標準 SQL クエリおよびJDBC APIが使用できるようになります。詳細については、「[15 分以内で理解できる Phoenix](#)」を参照してください。Phoenix は、Amazon EMR リリースバージョン 4.7.0 以降に含まれています。

古いバージョンの Amazon EMR から Amazon EMR リリースバージョン 5.4.0 以降にアップグレードして、セカンダリインデックスを使用する場合、[Apache Phoenix のドキュメント](#)で説明されるようにローカルインデックスをアップグレードしてください。Amazon EMR は必要な設定を hbase-site 分類から削除しますが、インデックスはデータを再入力する必要があります。インデックスはオンラインとオフラインでアップグレードできます。オンラインのアップグレードがデフォルトです。これはバージョン4.8.0以降のPhoenixクライアントで初期する間にインデックスの値が再設定されることを意味します。オフラインアップグレードを指定するには、`phoenix.client.localIndexUpgrade` 構成を `phoenix-site` 分類で `False` に設定してから、SSH をマスターノードに設定して `psql [zookeeper] -1`を実行します。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Phoenix のバージョンと、Amazon EMR で Phoenix と共にインストールされるコンポーネントを示しています。

このリリースで Phoenix と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Phoenix バージョン情報

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-7.1.0	Phoenix 5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
		timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Phoenix のバージョンと、Amazon EMR で Phoenix と共にインストールされるコンポーネントを示しています。

このリリースで Phoenix と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Phoenix バージョン情報

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.15.0	Phoenix 5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
		server, zookeeper-client, zookeeper-server

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Phoenix のバージョンと、Amazon EMR で Phoenix と共にインストールされるコンポーネントを示しています。

このリリースで Phoenix と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

#### emr-5.36.2 の Phoenix バージョン情報

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.36.2	Phoenix 4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

#### トピック

- [Phoenix を使用したクラスターの作成](#)
- [Phoenix クライアント](#)

- [Phoenix リリース履歴](#)

## Phoenix を使用したクラスターの作成

コンソールまたは AWS CLI でクラスターを作成するときにアプリケーションを選択することにより、Phoenix をインストールします。次の手順と例では、Phoenix と HBase を備えたクラスターを作成する方法を示します。詳細オプションなど、コンソールを使用したクラスター作成の詳細については、「Amazon EMR 管理ガイド」の「[クラスターの計画と構成](#)」を参照してください。

コンソールでクラスターを作成するために [Quick Options (クイックオプション)] を使用して、Phoenix がインストールされたクラスターを起動するには

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Create cluster (クラスターの作成)] を選択して [Quick Create (クイック作成)] を使用します。
3. [Software Configuration (ソフトウェア設定)] で、アプリケーションに適切な最近のリリースを選択します。[Amazon Release Version emr-4.7.0 (Amazon リリースバージョン emr-4.7.0)] 以降を選択した場合にのみ、Phoenix がオプションとして表示されます。
4. アプリケーションでは、2 番目のオプションとして、HBase : HBase **ver** with Ganglia **ver**、Hadoop **ver**、Hive **ver**、Hue **ver**、Phoenix **ver**、および ZooKeeper **ver** を選択します。
5. 必要に応じて他のオプションを選択し、[Create cluster (クラスターの作成)] を選択します。

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

次の例では、デフォルトの構成設定を使用して Phoenix がインストールされたクラスターを起動します。

を使用して Phoenix と HBase でクラスターを起動するには AWS CLI

- 次のコマンドを使用してクラスターを作成します。

```
aws emr create-cluster --name "Cluster with Phoenix" --release-label emr-7.1.0 \  
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \  
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

## Phoenix 設定のカスタマイズ

クラスターを作成するとき、hbase-site.xml 構成分類を使用して hbase-site 内で値を設定することによって Phoenix を構成します。

詳細については、Phoenix ドキュメントの「[Configuration and Tuning](#)」を参照してください。

次の例では、Amazon S3 に保存されている JSON ファイルを使用し、`phoenix.schema.dropMetaData` プロパティに `false` の値を指定しています。1つの分類に対して複数のプロパティを指定できます。詳細については、「[アプリケーションの設定](#)」を参照してください。create-cluster コマンドはその後、JSON ファイルを `--configurations` パラメータとして参照します。

/mybucket/myfolder/myconfig.json に保存される JSON ファイルのコンテンツは次のとおりです。

```
[  
  {  
    "Classification": "hbase-site",  
    "Properties": {  
      "phoenix.schema.dropMetaData": "false"  
    }  
  }  
]
```

JSON ファイルを参照する create cluster コマンドは、次の例に示されています。

```
aws emr create-cluster --release-label emr-7.1.0 --applications Name=Phoenix \  
Name=HBase --instance-type m5.xlarge --instance-count 2 \  
--configurations https://s3.amazonaws.com/mybucket/myfolder/myconfig.json
```

### Note

Phoenix の設定分類の再設定リクエストは、Amazon EMR バージョン 5.23.0 以降でのみサポートされており、Amazon EMR バージョン 5.21.0 または 5.22.0 ではサポートされてい

ません。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## Phoenix クライアント

Phoenix には、完全な依存関係によって構築された JDBC クライアント、または Phoenix クエリサーバーを使用しクラスターのマスターノードでのみ実行される「シンクライアント」を使用して接続します (例: SQL クライアント、ステップ、コマンドライン、SSH ポート転送などを使用)。「ファット」JDBC クライアントを使用している場合は、HBase サービスに直接接続するため、クラスターのすべてのノードにアクセスできる必要があります。「シン」Phoenix クライアントでは、Phoenix クエリサーバーにデフォルトポート 8765 でアクセスできる必要のみがあります。Phoenix には、これらのクライアントを使用する[スクリプト](#)がいくつかあります。

### Phoenix を使用したクエリでの Amazon EMR ステップの使用

次の手順は、HBase からスナップショット復元し、そのデータを使用して Phoenix クエリを実行します。必要に応じてこの例を拡張するか、Phoenix クライアントを利用する新しいスクリプトを作成できます。

1. 次のコマンドを使用して、Phoenix がインストールされた状態でクラスターを作成します。

```
aws emr create-cluster --name "Cluster with Phoenix" --log-uri s3://myBucket/myLogFolder --release-label emr-7.1.0 \  
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \  
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

2. 次のファイルを作成して、Amazon S3 にアップロードします。

#### copySnapshot.sh


```
sudo su hbase -s /bin/sh -c 'hbase snapshot export \  
-D hbase.rootdir=s3://us-east-1.elasticmapreduce.samples/hbase-demo-customer-data/snapshot/ \  
-snapshot customer_snapshot1 \  
-copy-to hdfs://masterDNSName:8020/user/hbase \  
-mappers 2 -chuser hbase -chmod 700'
```

#### runQuery.sh



```
aws s3 cp s3://myBucket/phoenixQuery.sql /home/hadoop/  
/usr/lib/Phoenix/bin/sqlline-thin.py http://localhost:8765 /home/hadoop/  
phoenixQuery.sql
```

phoenixQuery.sql

 Note

Amazon EMR versions 5.26.0 以降を使用する場合は、以下の例に COLUMN\_ENCODED\_BYTES=0 を含めるだけです。

```
CREATE VIEW "customer" (  
pk VARCHAR PRIMARY KEY,  
"address"."state" VARCHAR,  
"address"."street" VARCHAR,  
"address"."city" VARCHAR,  
"address"."zip" VARCHAR,  
"cc"."number" VARCHAR,  
"cc"."expire" VARCHAR,  
"cc"."type" VARCHAR,  
"contact"."phone" VARCHAR)  
COLUMN_ENCODED_BYTES=0;  
  
CREATE INDEX my_index ON "customer" ("customer"."state") INCLUDE("PK",  
"customer"."city", "customer"."expire", "customer"."type");  
  
SELECT "customer"."type" AS credit_card_type, count(*) AS num_customers FROM  
"customer" WHERE "customer"."state" = 'CA' GROUP BY "customer"."type";
```

を使用して AWS CLI S3 バケットにファイルを送信します。

```
aws s3 cp copySnapshot.sh s3://myBucket/  
aws s3 cp runQuery.sh s3://myBucket/  
aws s3 cp phoenixQuery.sql s3://myBucket/
```

3. ステップ 1 で作成したクラスターに送信された以下の手順を使用してテーブルを作成します。

createTable.json

```
[
  {
    "Name": "Create HBase Table",
    "Args": ["bash", "-c", "echo '$create \"customer\", \"address\", \"cc\", \"contact\" | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file:///./createTable.json
```

4. `script-runner.jar` を使用して、先に S3 バケットにアップロードした `copySnapshot.sh` スクリプトを実行します。

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="HBase Copy Snapshot",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://myBucket/copySnapshot.sh"]
```

これにより、MapReduce ジョブが実行され、スナップショットデータがクラスター HDFS にコピーされます。

5. 次のステップを使用して、クラスターにコピーしたスナップショットを復元します。

`restoreSnapshot.json`

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo '$disable \"customer\"; restore_snapshot \"customer_snapshot1\"; enable \"customer\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file:///./restoreSnapshot.json
```

6. `script-runner.jar` を使用して、先に S3 バケットにアップロードした `runQuery.sh` スクリプトを実行します。

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Phoenix Run Query",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://myBucket/runQuery.sh"]
```

クエリが実行され、ステップの `stdout` に結果が返ります。このステップが完了するには数分間かかる場合があります。

7. ステップ 1 でクラスターを作成したときに使用したログ URI でステップの `stdout` の結果を確認します。結果は次のようになります。

```
+-----+-----+
|          CREDIT_CARD_TYPE          |          NUM_CUSTOMERS          |
+-----+-----+
| american_express                    | 5728                             |
| dankort                              | 5782                             |
| diners_club                          | 5795                             |
| discover                             | 5715                             |
| forbrugsforeningen                  | 5691                             |
| jcb                                  | 5762                             |
| laser                               | 5769                             |
| maestro                             | 5816                             |
| mastercard                          | 5697                             |
| solo                                 | 5586                             |
| switch                              | 5781                             |
| visa                                 | 5659                             |
+-----+-----+
```

## Phoenix リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Phoenix のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または

[Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

## Phoenix バージョン情報

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.36.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-7.1.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tool

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
		ls, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server
emr-7.0.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.15.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.14.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.13.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.12.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.11.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.11.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.10.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.10.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.9.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.9.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.8.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.8.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.7.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.36.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.36.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.6.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.35.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.5.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.4.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.3.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.3.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.2.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.2.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.1.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.1.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.0.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-6.0.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.34.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.33.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.33.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.32.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.32.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.31.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.31.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.30.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.30.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.30.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.29.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.28.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.28.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.27.1	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.27.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.26.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.25.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.24.1	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.24.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.23.1	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.23.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.22.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.21.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.20.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.20.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.19.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.19.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.18.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.18.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.17.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.17.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.17.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.16.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.16.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.15.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.15.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.14.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.13.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.13.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.12.3	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.12.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.11.4	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.11.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.10.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.10.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.9.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.8.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.7.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.6.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.6.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.4	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.5.3	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.2	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.5.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.4.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.3.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.3.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.3.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.2.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.2.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.1.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.1.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.0.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-5.0.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.6	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-4.9.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-4.9.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-4.9.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-4.8.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-4.8.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-4.8.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-4.7.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	Phoenix バージョン	Phoenix でインストールされるコンポーネント
emr-4.7.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpserver, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-master, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

# Apache Pig

Apache Pig はオープンソースの Apache ライブラリで、Hadoop 上で稼動し、大規模なデータセットを複雑なコードを記述することなく Java などの低レベルコンピュータ言語で変換するのに使用できるスクリプト言語を提供します。このライブラリは、Pig Latin という言語で記述された SQL のようなコマンドを受け取り、それらのコマンドを有向非巡回グラフ (DAGs) ジョブに変換します。MapReduce Pig は多様な形式の構造化データと非構造化データで機能します。Pig の詳細については、<http://pig.apache.org/> を参照してください。

Pig コマンドは、インタラクティブに実行するか、バッチモードで実行できます。Pig をインタラクティブに使用するには、マスターノードへの SSH 接続を作成し、Grunt シェルを使用してコマンドを送信します。Pig をバッチモードで使用するには、Pig スクリプトを記述し、Amazon S3 にアップロードして、クラスターステップとして送信します。クラスターへの作業の送信の詳細については、「Amazon EMR 管理ガイド」の「[クラスターへの作業の送信](#)」を参照してください。

Pig を使用して Amazon S3 の HCatalog テーブルに出力を書き込む場合は、`mapred.output.direct.NativeS3FileSystem` プロパティと `mapred.output.direct.EmrFileSystem` プロパティを `false` に設定して、Amazon EMR の直接書き込み機能を無効にしてください。詳細については、「[HCatalog の使用](#)」を参照してください。Pig スクリプト内で、`SET mapred.output.direct.NativeS3FileSystem false` および `SET mapred.output.direct.EmrFileSystem false` コマンドを使用できます。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Pig のバージョンと、Amazon EMR で Pig と共にインストールされるコンポーネントを示しています。

このリリースで Pig と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Pig バージョン情報

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-7.1.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
		hdfs-namenode, hadoop- httpfs-server, hadoop-kms- server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn- timeline-server, pig-client, tez- on-yarn, tez-on-worker

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Pig のバージョンと、Amazon EMR で Pig と共にインストールされるコンポーネントを示しています。

このリリースで Pig と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Pig バージョン情報

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.15.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodi es, emr-kinesis, emr-s3-dist- cp, hadoop-client, hadoop-ma pred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop- hdfs-namenode, hadoop- httpfs-server, hadoop-kms- server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn- timeline-server, pig-client, tez- on-yarn, tez-on-worker

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Pig のバージョンと、Amazon EMR で Pig と共にインストールされるコンポーネントを示しています。

このリリースで Pig と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

### emr-5.36.2 の Pig バージョン情報

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.36.2	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

### トピック

- [Pig の作業の送信](#)
- [Pig からユーザー定義関数を呼び出す](#)
- [Pig リリース履歴](#)

## Pig の作業の送信

このセクションでは、Pig の作業を Amazon EMR クラスターに送信する方法を示します。次の例では、Apache ログファイルを評価し、転送された合計バイト数、上位 50 の IP アドレスリスト、上位 50 の外部参照者リスト、Bing や Google を使用した検索語トップ 50 などの情報が含まれるレポートを生成します。Pig スクリプトは、Amazon S3 バケット `s3://elasticmapreduce/samples/pig-apache/do-reports2.pig` に配置されています。入力データは Amazon S3 バケット `s3://elasticmapreduce/samples/pig-apache/input` に配置されています。出力は、Amazon S3 バケットに保存されます。

## Amazon EMR コンソールを使用した Pig の作業の送信

この例では、Amazon EMR コンソールを使用し、Pig のステップをクラスターに追加する方法について説明します。

Pig のステップを送信するには

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Create cluster] (クラスターの作成) を選択し、Pig がインストールされたクラスターを作成します。クラスターの作成手順については、「[Amazon EMR クラスターを計画して構成する](#)」を参照してください。
3. ターミナルを開き、「[SSH を使用してマスターノードに接続する](#)」で説明されている手順に従ってクラスターのマスターノードに SSH 接続します。完了したら、次の手順を実行してください。

```
sudo mkdir -p /home/hadoop/lib/pig/  
sudo aws s3 cp s3://elasticmapreduce/libs/pig/0.3/piggybank-0.3-amzn.jar /home/  
hadoop/lib/pig/piggybank.jar
```

4. コンソールで [Cluster List] (クラスターリスト) をクリックし、作成したクラスターの名前を選択します。
5. [Steps (ステップ)] セクションまでスクロールして展開し、[Add step (ステップの追加)] を選択します。
6. [Add Step (ステップの追加)] ダイアログで:
  - [Step type (ステップタイプ)] で、[Pig program (Pig プログラム)] を選択します。
  - [Name (名前)] では、デフォルト名 ( Pig program ) を使用するか、または新しい名前を入力します。
  - [Script S3 location (スクリプト S3 の場所)] で、Pig スクリプトの場所を入力します。例: **s3://elasticmapreduce/samples/pig-apache/do-reports2.pig**。
  - [Input S3 location (S3 の場所の入力)] で、入力データの場所を入力します。例: **s3://elasticmapreduce/samples/pig-apache/input**。
  - [Output S3 location] (S3 の場所の出力) で、Amazon S3 出力バケットの名前を入力するか、参照します。
  - [Arguments (引数)] は、フィールドを空のままにします。



- [Action on failure (失敗時の操作)] では、デフォルトのオプション ([Continue (続行)]) を使用します。
7. [追加] を選択します。ステップは、[Pending] というステータスでコンソールに表示されます。
  8. ステップが実行されると、ステータスは [Pending (保留中)] から [Running (実行中)]、[Running (実行中)] から [完了済み] に変更されます。ステータスを更新するには、[Actions] (アクション) 列の上にある [Refresh] (更新) アイコンを選択します。ステップが完了したら、Amazon S3 バケットをチェックして Pig ステップの出カファイルがあることを確認します。

## を使用して Pig 作業を送信する AWS CLI

を使用して Pig ステップを送信するには AWS CLI

を使用してクラスターを起動するときは AWS CLI、`--applications` パラメータを使用して Pig をインストールします。Pig のステップを送信するには、`--steps` パラメータを使用します。

1. Pig がインストールされたクラスターを起動するには、次のコマンドを入力します。*myKey* と *DOC-EXAMPLE-BUCKET/* は、EC2 キーペアと Amazon S3 バケットの名前に置き換えます。

```
aws emr create-cluster \  
--name "Test cluster" \  
--log-uri s3://DOC-EXAMPLE-BUCKET/ \  
--release-label emr-5.36.2 \  
--applications Name=Pig \  
--use-default-roles \  
--ec2-attributes KeyName=myKey \  
--instance-type m5.xlarge \  
--instance-count 3
```

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

`--instance-groups` パラメータを使用せずにインスタンス数を指定すると、1 つのマスターノードが起動され、残りのインスタンスはコアノードとして起動されます。すべてのノードで、コマンドで指定したインスタンスタイプが使用されます。

**Note**

以前にデフォルトの EMR サービスロールと EC2 インスタンスプロファイルを作成していない場合は、「aws emr create-default-roles」と入力してそれらを作成してから、create-cluster サブコマンドを入力します。

2. Pig ステップを送信するには、次のコマンドを入力し、*myClusterId*と *DOC-EXAMPLE-BUCKET* をクラスター ID と Amazon S3 バケットの名前に置き換えます。

```
aws emr add-steps \  
--cluster-id myClusterId \  
--steps Type=PIG,Name="Pig Program",ActionOnFailure=CONTINUE,Args=[-f,s3://  
elasticmapreduce/samples/pig-apache/do-reports2.pig, -p,INPUT=s3://elasticmapreduce/  
samples/pig-apache/input, -p,OUTPUT=s3://DOC-EXAMPLE-BUCKET/pig-apache/output]
```

このコマンドを実行すると、ステップ ID が返されます。これを使用して、ステップの State を確認できます。

3. describe-step コマンドを使用して、ステップのステータスのクエリを実行します。

```
aws emr describe-step --cluster-id myClusterId --step-id s-1XXXXXXXXXXA
```

ステップの State は、ステップの実行に伴って PENDING から RUNNING、さらには COMPLETED へと変わります。ステップが完了したら、Amazon S3 バケットをチェックして Pig ステップの出カファイルがあることを確認します。

での Amazon EMR コマンドの使用の詳細については AWS CLI、[AWS CLI 「コマンドリファレンス」](#)を参照してください。

## Pig からユーザー定義関数を呼び出す

Pig では、ユーザー定義関数 (UDF) を Pig スクリプトから呼び出すことができます。これにより、Pig スクリプト内で使用するカスタム処理を実装できます。現在サポートされている言語は Java、Python/Jython、および JavaScript (ただし、サポートはまだ実験的です JavaScript) です。

以下のセクションでは、関数を Pig シェルまたは Pig スクリプトから呼び出せるように、Pig に登録する方法を説明します。Pig での UDF の使用の詳細については、ご使用のバージョンの「[Pig のドキュメント](#)」を参照してください。

## Pig から JAR ファイルを呼び出す

Pig スクリプトに REGISTER コマンドを使用して、Pig とともにカスタム JAR ファイルを使用できます。JAR ファイルは、ローカルファイルシステム、または Amazon S3 などのリモートファイルシステムです。Pig スクリプトが実行されると、Amazon EMR は自動的に JAR ファイルをマスターノードにダウンロードし、その JAR ファイルを Hadoop 分散キャッシュにアップロードします。これにより、JAR ファイルは必要に応じてクラスター内のすべてのインスタンスで自動的に使用されます。

Pig で JAR ファイルを使用するには

1. Amazon S3 にカスタム JAR ファイルをアップロードします。
2. Pig スクリプト内で REGISTER コマンドを使用して、カスタム JAR ファイルの Amazon S3 バケットを指定します。

```
REGISTER s3://mybucket/path/mycustomjar.jar;
```

## Pig から Python/Jython スクリプトを呼び出す

Python スクリプトを Pig に登録し、登録したスクリプト内の関数を Pig シェルまたは Pig スクリプトから呼び出すことができます。これを行うには、`register` キーワードでスクリプトの場所を指定します。

Pig は Java で記述されているため、Python スクリプトの解析には Jython スクリプトエンジンが使用されます。Jython の詳細については、<http://www.jython.org/> を参照してください。

Pig から Python/Jython スクリプトを呼び出すには

1. Python スクリプトを作成して、Amazon S3 内の場所にアップロードします。この場所は、Pig クラスターを作成するアカウントと同じアカウントが所有しているか、クラスターを作成したアカウントがアクセスできるように許可が設定されているバケットである必要があります。この例では、スクリプトは `s3://mybucket/pig/python` にアップロードされます。
2. Pig クラスターを起動します。Grunt シェルから Pig にアクセスする場合は、インタラクティブクラスターを起動します。Pig コマンドをスクリプトから実行している場合は、スクリプトで記述された Pig クラスターを起動します。この例では、インタラクティブクラスターを起動します。Pig クラスターの詳細については、「[Pig の作業の送信](#)」を参照してください。

3. インタラクティブクラスターの場合は、SSH を使用してマスターノードに接続し、Grunt シェルを実行します。詳細については、「[マスターノードへの SSH 接続](#)」を参照してください。
4. コマンドラインに pig と入力することにより、Pig 向けに Grunt を実行します。

```
pig
```

5. 以下のコマンドに示すように Grunt コマンドプロンプトで register キーワードを使用し、Amazon S3 内のスクリプトの場所を指定して Jython ライブラリおよび Python スクリプトを Pig に登録します。

```
grunt> register 'lib/jython.jar';
grunt> register 's3://mybucket/pig/python/myscript.py' using jython as myfunctions;
```

6. 入力データをロードします。次の例では、入力を Amazon S3 上の場所からロードします。

```
grunt> input = load 's3://mybucket/input/data.txt' using TextLoader as
(line:chararray);
```

7. これにより、myfunctions を使用してスクリプト内の関数を参照して、Pig から呼び出すことができます。

```
grunt> output=foreach input generate myfunctions.myfunction($1);
```

## Pig リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Pig のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Pig バージョン情報

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.36.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-ma

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
		pred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-7.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-7.0.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker



Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.8.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.36.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.36.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.35.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.3.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn



Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.2.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-6.1.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.34.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.33.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.33.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.32.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.32.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.31.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.31.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.30.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.30.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.30.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.29.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.28.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn



Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.28.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.27.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.27.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.26.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.25.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.24.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.24.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.23.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.23.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.22.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.21.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.21.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.21.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.20.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.20.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.19.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn



Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.19.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.18.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.18.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.17.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.17.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.17.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.16.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.16.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.15.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.14.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.14.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.13.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn



Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.12.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.4	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.11.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.8.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.8.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn



Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.7.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.7.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.6.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.6.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.5.4	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.5.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.5.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.4.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.4.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.3.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.3.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.3.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.2.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn



Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.2.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.1.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.1.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.0.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-5.0.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.9.6	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.9.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.9.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.9.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.9.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.9.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.8.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.8.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.8.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.8.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.8.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.8.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client



Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.7.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.7.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client
emr-4.7.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, pig-client

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.7.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.6.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.6.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.5.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.4.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.3.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.2.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.1.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httplibfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR リリースラベル	Pig のバージョン	Pig でインストールされるコンポーネント
emr-4.0.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datano-de, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

# Presto と Trino

## Note

2020 年 12 月に、PrestoSQL は Trino に名前変更されました。Amazon EMR バージョン 6.4.0 以降では Trino という名称が使用されていますが、それ以前のリリースバージョンでは PrestoSQL という名称が使用されています。

[Presto](#) は、複数のソースからの大量のデータセットの対話型分析クエリ用に設計された高速の SQL クエリエンジンです。詳細については、[Presto のウェブサイト](#)を参照してください。Presto は、Amazon EMR リリース 5.0.0 以降に含まれています。以前のリリースには、サンドボックスアプリケーションとして Presto が含まれています。詳細については、「[Amazon EMR 4.x リリース バージョン](#)」を参照してください。Amazon EMR リリースバージョン 6.1.0 以降では、Presto に加えて [Trino](#) (PrestoSQL) がサポートされています。詳細については、「[PrestoDB と Trino のインストール](#)」を参照してください。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Presto のバージョンと、Amazon EMR で Presto と共にインストールされるコンポーネントを示しています。

このリリースで Presto と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Presto バージョン情報

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-7.1.0	Presto 0.284	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto,

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
		hcatalog-server, mariadb-server, presto-coordinator, presto-worker

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Presto のバージョンと、Amazon EMR で Presto と共にインストールされるコンポーネントを示しています。

このリリースで Presto と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Presto バージョン情報

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.15.0	Presto 0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Presto のバージョンと、Amazon EMR で Presto と共にインストールされるコンポーネントを示しています。

このリリースで Presto と共にインストールされるコンポーネントのバージョンについては、「[リリース 5.36.2 コンポーネントバージョン](#)」を参照してください。

## emr-5.36.2 の Presto バージョン情報

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.36.2	Presto 0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Trino (PrestoSQL) のバージョンと、Amazon EMR で Trino (PrestoSQL) と共にインストールされるコンポーネントを示しています。

このリリースで Trino (PrestoSQL) と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の Trino (PrestoSQL) バージョン情報

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.15.0	Trino (PrestoSQL) 426	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client,



Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
		hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

## トピック

- [AWS Glue データカタログでの Presto の使用](#)
- [S3 Select Pushdown と Presto を使用してパフォーマンスを向上させる](#)
- [データベースコネクタの追加](#)
- [Amazon EMR での SSL/TLS の使用と Presto による LDAPS の設定](#)
- [Presto 厳格モードを有効にする](#)
- [Presto でのスポットインスタンス損失の処理](#)
- [Trino での耐障害性実行](#)
- [グレースフルな廃止による Presto Auto Scaling の使用](#)
- [Amazon EMR での Presto に関する考慮事項](#)
- [Presto リリース履歴](#)

## AWS Glue データカタログでの Presto の使用

Amazon EMR リリースバージョン 5.10.0 以降を使用して、AWS Glue データカタログを Presto のデフォルトの Hive メタストアとして指定できます。永続的なメタストア、またはさまざまなクラスター、サービス、アプリケーション、あるいは AWS アカウントで共有されるメタストアが必要な場合は、この設定をお勧めします。

AWS Glue はフルマネージド型の抽出、変換、ロード (ETL) サービスで、データの分類、クリーニング、強化、さまざまなデータストア間の信頼性の高い移動を簡単かつ費用対効果の高い方法で実現します。AWS Glue Data Catalog は、Amazon EMR だけでなく、Amazon RDS、Amazon Redshift、Redshift Spectrum、Athena、および Apache Hive メタストアと互換性のあるすべてのアプリケーションと統合されたメタデータリポジトリを提供します。AWS Glue クローラは、Amazon S3 のソースデータからスキーマを自動的に推測し、関連するメタデータをデータカタログに保存できます。データカタログの詳細については、「[AWS Glue デベロッパーガイド](#)」の「[Glue データカタログの入力](#)」を参照してください。AWS

AWS Glue には個別の料金が適用されます。Data Catalog にメタデータを保存してアクセスするための月額料金、AWS Glue ETL ジョブとクローラランタイムに対して 1 分あたりに課金される時間料金、プロビジョニングされた開発エンドポイントごとに 1 分あたりに課金される時間料金があります。Data Catalog では、最大 100 万個までのオブジェクトを無料で保存できます。100 万を超えるオブジェクトを保存した場合は、100,000 オブジェクトごとに 1 USD が課金されます。Data Catalog 内のオブジェクトは、テーブル、パーティション、またはデータベースです。詳細については、「[Glue 料金表](#)」を参照してください。

#### Important

2017 年 8 月 14 日より前に Amazon Athena または Amazon Redshift Spectrum を使用してテーブルを作成した場合、データベースとテーブルは AWS Glue データカタログとは別の Athena 管理カタログに保存されます。Amazon EMR をこれらのテーブルと統合するには、AWS Glue データカタログにアップグレードする必要があります。詳細については、「[Amazon Athena ユーザーガイド](#)」の「[AWS Glue データカタログへのアップグレード](#)」を参照してください。Amazon Athena

## AWS Glue Data Catalog をメタストアとして指定する

、または Amazon EMR API を使用して AWS Management Console AWS CLI、AWS Glue データカタログをメタストアとして指定できます。CLI または API を使用する場合は、Presto の設定分類を使用してデータカタログを指定します。さらに、Amazon EMR 5.16.0 以降では、設定分類を使用して、別のデータカタログを指定できます AWS アカウント。コンソールを使用する場合は、[Advanced Options] (詳細オプション) または [Quick Options] (クイックオプション) を使用して、Data Catalog を指定できます。

### New console

新しいコンソールで AWS Glue Data Catalog を Hive メタストアとして指定するには

1. にサインインし AWS Management Console、<https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. 左側のナビゲーションペインの [EMR on EC2] で、[クラスター] を選択し、[クラスターの作成] を選択します
3. [アプリケーションバンドル] で [Presto] を選択します。
4. [AWS Glue Data Catalog 設定] で [Presto テーブルメタデータに使用] チェックボックスをオンにします。

5. クラスターに適用するその他のオプションを選択します。
6. クラスターを起動するには、[クラスターの作成] を選択します。

## Old console

古いコンソールで AWS Glue データカタログをデフォルトの Presto メタストアとして指定するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Software Configuration (ソフトウェア設定)] の [Release (リリース)] で、[emr-5.10-0] 以降を選択し、続いて [Presto] を選択します。
4. [Presto テーブルメタデータに使用]、[Next (次へ)] の順に選択し、アプリケーションに該当するクラスターのその他の設定を行います。

## CLI

を使用して AWS Glue データカタログをデフォルトの Hive メタストアとして指定するには AWS CLI

クラスターの作成時に以下の設定分類を指定する方法の例については、「[アプリケーションの設定](#)」を参照してください。

### Amazon EMR 5.16.0 以降

- 以下の JSON の例に示すように、hive.metastore プロパティを glue に設定します。

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
]
```

別のデータカタログを指定するには AWS アカウント、次の JSON の例に示すように `hive.metastore.glue.catalogid` プロパティを追加します。`acct-id` は、データカタログの AWS アカウントに置き換えます。Amazon EMR バージョン 5.15.0 以前では、別の AWS アカウントのデータカタログは使用できません。

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

#### Amazon EMR 5.10.0 ~ 5.15.0

以下の JSON の例に示すように、`hive.metastore.glue.datacatalog.enabled` プロパティを `true` に設定します。

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore.glue.datacatalog.enabled": "true"
    }
  }
]
```

#### Amazon EMR 6.1.0 以降で PrestoSQL (Trino) を使用

EMR バージョン 6.1.0 以降、PrestoSQL は Glue もデフォルトの Hive メタストアとしてサポートしています。以下の JSON の例に示されているように、`prestoql-connector-hive` 設定分類を使用し、`hive.metastore` プロパティを `glue` に設定します。

Amazon EMR バージョン 6.4.0 以降では、PrestoSQL ではなく、新しい名称の Trino を使用します。Trino を使用する場合、以下の設定分類で `prestoql-connector-hive` を `trino-connector-hive` に置き換えてください。

```
[
  {
    "Classification": "prestoql-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
]
```

長時間稼働クラスターでメタストアを切り替えるには、マスターノードに接続して `/etc/presto/conf/catalog/hive.properties` ファイルのプロパティ値を直接編集し、Presto サーバー (`sudo restart presto-server`) を再起動することにより、リリースバージョンに最適な値を手動で設定できます。この方法を Amazon EMR 5.15.0 以前のバージョンで使用する場合は、`hive.table-statistics-enabled` が `false` に設定されていることを確認します。リリースバージョン 5.16.0 以降を使用する場合、この設定は必要ありません。ただし、テーブルとパーティションの統計はサポートされません。

## IAM アクセス許可

クラスターの EC2 インスタンスプロファイルには、AWS Glue アクションの IAM アクセス許可が必要です。さらに、AWS Glue Data Catalog オブジェクトの暗号化を有効にする場合は、暗号化 AWS KMS key に使用される の暗号化、復号、生成もロールに許可する必要があります。

### AWS Glue アクションのアクセス許可

Amazon EMR のデフォルトの EC2 インスタンスプロファイルを使用する場合、アクションは必要ありません。にアタッチされている `AmazonElasticMapReduceforEC2Role` マネージドポリシーは、必要なすべての AWS Glue アクション `EMR_EC2_DefaultRole` を許可します。ただし、カスタム EC2 インスタンスプロファイルとアクセス許可を指定する場合は、適切な AWS Glue アクションを設定する必要があります。開始点として `AmazonElasticMapReduceforEC2Role` 管理ポリシーを使用します。詳細については、「Amazon EMR 管理ガイド」の「[クラスター EC2 インスタンスのサービスロール \(EC2 インスタンスプロファイル\)](#)」を参照してください。

## AWS Glue Data Catalog を暗号化および復号するためのアクセス許可

インスタンスプロファイルには、キーを使用してデータを暗号化および復号するためのアクセス許可が必要です。以下のステートメントが両方とも適用される場合、これらのアクセス許可を設定する必要はありません。

- AWS Glue の マネージドキーを使用して、AWS Glue Data Catalog オブジェクトの暗号化を有効にします。
- AWS Glue データカタログ AWS アカウント と同じ にあるクラスターを使用します。

それ以外の場合は、EC2 インスタンスプロファイルにアタッチされたアクセス許可ポリシーに次のステートメントを追加する必要があります。

```
[
  {
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Action": [
          "kms:Decrypt",
          "kms:Encrypt",
          "kms:GenerateDataKey"
        ],
        "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
      }
    ]
  }
]
```

AWS Glue データカタログの暗号化の詳細については、「[AWS Glue デベロッパーガイド](#)」の「[データカタログの暗号化](#)」を参照してください。

## リソースベースのアクセス許可

Amazon EMR で AWS Glue を Hive、Spark、または Presto と組み合わせて使用する場合、AWS Glue は Data Catalog リソースへのアクセスを制御するためのリソースベースのポリシーをサポートします。これらのリソースには、データベース、テーブル、接続、ユーザー定義関数が含まれます。

詳細については、「AWS Glue デベロッパーガイド」の「[AWS Glue リソースポリシー](#)」を参照してください。

リソースベースのポリシーを使用して Amazon EMR 内から AWS Glue へのアクセスを制限する場合、アクセス許可ポリシーで指定するプリンシパルは、クラスターの作成時に指定する EC2 インスタンスプロファイルに関連付けられたロール ARN である必要があります。例えば、カタログにアタッチされたリソースベースのポリシーの場合、次の例に示す形式を使用して、クラスター EC2 インスタンスのデフォルトのサービスロールのロール ARN `Principal`、`EMR_EC2_DefaultRole` をとして指定できます。

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

`acct-id` は AWS Glue アカウント ID とは異なる場合があります。これにより、さまざまなアカウントで EMR クラスターからアクセスできます。異なるアカウントから、複数のプリンシパルを指定できます。

## AWS Glue Data Catalog を使用する場合の考慮事項

Presto でメタストアとして AWS Glue Data Catalog を使用する場合は、次の項目を考慮してください。

- AWS Glue 内からのテーブルの名前変更はサポートされていません。
- LOCATION を指定せずに Hive テーブルを作成すると、テーブルデータは、`hive.metastore.warehouse.dir` プロパティによって指定された場所に保管されます。デフォルトでは、これは HDFS 内の場所です。別のクラスターがテーブルにアクセスする必要がある場合、テーブルを作成したクラスターに対する適切なアクセス許可がない限り、処理に失敗します。さらに、HDFS ストレージは一時的であるため、クラスターが終了すると、テーブルデータは失われ、テーブルを再作成する必要があります。AWS Glue を使用して Hive Amazon S3 LOCATION で指定することをお勧めします。または、`hive-site` 設定分類を使用して、`hive.metastore.warehouse.dir` の Amazon S3 で場所を指定できます。これは、すべての Hive テーブルに適用されます。テーブルが HDFS ロケーションで作成され、テーブルを作成したクラスターがまだ実行されている場合は、AWS Glue 内からテーブルロケーションを Amazon S3 に更新できます。詳細については、「[AWS Glue デベロッパーガイド](#)」の「[Glue コンソールでのテーブルの操作](#)」を参照してください。AWS
- 引用符とアポストロフィを含むパーティション値はサポートされていません(例: `PARTITION (owner="Doe 's")`.)。
- [列統計](#)は、`emr-5.31.0` 以降でサポートされています。



- [Hive 認可](#)の使用はサポートされていません。代替策として、[AWS Glue リソースベースのポリシー](#)を使用することを検討してください。詳細については、[AWS 「Glue データカタログへの Amazon EMR アクセスにリソースベースのポリシーを使用する」](#)を参照してください。

## S3 Select Pushdown と Presto を使用してパフォーマンスを向上させる

Amazon EMR リリースバージョン 5.18.0 以降では、Amazon EMR での Presto を使用した [S3 select](#) プッシュダウンを使用できます。この機能により、Presto から Amazon S3 に (SELECT などの) 射影演算と (WHERE などの) 述語演算の計算処理を「プッシュダウン」できます。これにより、一部のアプリケーションではクエリで Amazon S3 から必要なデータだけを取得してパフォーマンスを向上させ、Amazon EMR と Amazon S3 間で転送されるデータの量を減らすことができます。

### S3 Select Pushdown が使用するアプリケーションに適しているかどうかを確認する

S3 Select Pushdown が使用するアプリケーションに適しているかどうかを確認するために、S3 Select Pushdown を使用した場合と使用しない場合のアプリケーションのベンチマークを行うことをお勧めします。

アプリケーションが S3 Select を使用する候補となるかどうかを判断するには、次のガイドラインを使用します。

- クエリは元のデータセットの半分以上を除外する。
- クエリのフィルター述語では、Presto と S3 Select がサポートするデータ型を持つ列が使用されます。S3 Select Pushdown では、タイムスタンプデータ型、実データ型、および倍精度浮動小数点データ型はサポートされません。数値データには小数データ型を使用することをお勧めします。S3 Select でサポートされているデータ型の詳細については、「Amazon Simple Storage Service ユーザーガイド」の「[データ型](#)」を参照してください。
- Amazon S3 と Amazon EMR クラスター間のネットワーク接続は、転送速度と使用可能な帯域幅が良好です。Amazon S3 は、HTTP 応答を圧縮しないため、応答サイズは圧縮された入力ファイルと比較して増加する可能性があります。



## 考慮事項と制約事項

- CSV 形式で保存されたオブジェクトのみがサポートされます。オブジェクトは解凍するか、必要に応じて gzip または bzip2 で圧縮できます。
- AllowQuotedRecordDelimiters プロパティはサポートされていません。このプロパティを指定した場合、クエリは失敗します。
- お客様が用意した暗号化キーを使用した Amazon S3 サーバー側の暗号化 (SSE-C) とクライアント側の暗号化はサポートされていません。
- S3 Select Pushdown は、ORC や Parquet などの列または圧縮ファイル形式には適していません。

## PrestoDB または Trino で S3 Select Pushdown を有効にする

Amazon EMR で PrestoDB の S3 Select Pushdown を有効にするには、以下の例に示されているように、`presto-connector-hive` 設定分類を使用して `hive.s3select-pushdown.enabled` を `true` に設定します。詳細については、「[アプリケーションの設定](#)」を参照してください。`hive.s3select-pushdown.max-connections` 値も設定する必要があります。ほとんどのアプリケーションでは、デフォルト設定の `500` で十分です。詳細については、以下の [hive.s3select-pushdown.max-connections の理解と調整](#) を参照してください。

EMR バージョン 6.1.0 - 6.3.0 での PrestoSQL の場合、以下の例で `presto-connector-hive` を `prestoql-connector-hive` に置き換えます。

Amazon EMR バージョン 6.4.0 以降では、PrestoSQL ではなく、新しい名称の Trino を使用します。Trino を使用する場合、以下の例の `presto-connector-hive` を `trino-connector-hive` に置き換えます

```
[
  {
    "classification": "presto-connector-hive",
    "properties": {
      "hive.s3select-pushdown.enabled": "true",
      "hive.s3select-pushdown.max-connections": "500"
    }
  }
]
```

## hive.s3select-pushdown.max-connections の理解と調整

デフォルトでは、Presto はファイルシステムとして EMRFS を使用します。emrfs-site 設定分類の設定 fs.s3.maxConnections により、Presto の EMRFS を使用した Amazon S3 へのクライアント接続の最大許容数が指定されます。デフォルトでは 500 です。述語演算のために Amazon S3 にアクセスすると、S3 Select Pushdown は EMRFS をバイパスします。この場合、hive.s3select-pushdown.max-connections の値によってワーカーノードからのオペレーションに使用可能なクライアント接続の最大数が決まります。ただし、(GET オペレーションなどの) プッシュダウンされていない Presto から Amazon S3 へのリクエストは、引き続き fs.s3.maxConnections の値に従います。

アプリケーションでエラー「Timeout waiting for connection from pool」が発生する場合は、hive.s3select-pushdown.max-connections と fs.s3.maxConnections 両方の値を増やします。

## データベースコネクタの追加

クラスターの作成時に設定分類を使用して JDBC コネクタのプロパティを設定できます。設定分類は presto-connector で始まります (例: presto-connector-postgresql)。Amazon EMR リリースバージョンによって使用可能な設定分類は異なります。最新のリリースバージョンで使用できる設定分類については、「for Amazon EMR 5.36.2 [the section called "5.36.2 設定分類"](#)」を参照してください。Amazon EMR の別のバージョンを使用している場合は、「[Amazon EMR 5.x リリースバージョン](#)」の設定分類を参照してください。各コネクタで設定できるプロパティの詳細については、<https://prestodb.io/docs/current/connector.html> を参照してください。

### Example PostgreSQL JDBC コネクタでのクラスターの設定

PostgreSQL コネクタがインストールおよび設定されたクラスターを起動するには、(myConfig.json などの) 設定分類を指定する、以下の内容が含まれた JSON ファイルを作成してからローカルに保存します。

Presto ドキュメントの「[PostgreSQL Connector](#)」トピックに示されているように、必要に応じてセットアップの接続プロパティを置き換えます。

```
[
  {
    "Classification": "presto-connector-postgresql",
    "Properties": {
```

```
    "connection-url": "jdbc:postgresql://example.net:5432/database",
    "connection-user": "MYUSER",
    "connection-password": "MYPASS"
  },
  "Configurations": []
}
]
```

クラスターを作成するときは、以下の例に示すように `--configurations` オプションを使用して JSON ファイルへのパスを参照します。myConfig.json は、コマンドを実行したのと同じディレクトリにあります。

```
aws emr create-cluster --name PrestoConnector --release-label emr-5.36.2 --instance-type m5.xlarge \
--instance-count 2 --applications Name=Hadoop Name=Hive Name=Pig Name=Presto \
--use-default-roles --ec2-attributes KeyName=myKey \
--log-uri s3://my-bucket/logs --enable-debugging \
--configurations file://myConfig.json
```

## Amazon EMR での SSL/TLS の使用と Presto による LDAPS の設定

Amazon EMR リリースバージョン 5.6.0 以降では、Presto ノード間の [セキュアな内部通信](#) のために SSL/TLS を有効にできます。これを行うには、転送時の暗号化のセキュリティ設定を指定します。詳細については、「Amazon EMR 管理ガイド」の「[暗号化オプション](#)」および「[セキュリティ設定を使用してクラスターセキュリティをセットアップする](#)」を参照してください。

セキュリティ設定を転送時の暗号化に使用すると、Amazon EMR は Presto に対して以下を実行します。

- 転送時の暗号化に指定した暗号化アーティファクトまたは証明書を、Presto クラスター全体に分散します。詳細については、「[伝送中のデータの暗号化のための証明書を提供する](#)」を参照してください。
- presto-config 設定分類を使用して、次のプロパティを設定します。これは Presto の config.properties ファイルに対応します。
  - すべてのノードで http-server.http.enabled を false に設定します。これにより、HTTP を無効にして HTTPS を使用します。これには、伝送中の暗号化のセキュリティ構成を設定するときに、パブリック DNS とプライベート DNS で動作する証明書を提供する必要があります。

あります。これを行う方法の1つは、複数のドメインをサポートする SAN (サブジェクトの別名) 証明書を使用することです。

- `http-server.https.*` 値を設定します。設定の詳細については、Presto ドキュメントの「[LDAP Authentication](#)」を参照してください。
- EMR バージョン 6.1.0 以降の PrestoSQL (Trino) では、Amazon EMR はクラスターノード間のセキュアな内部通信用に共有秘密キーを自動的に設定します。このセキュリティ機能を有効にするために追加の設定を行う必要はなく、独自の秘密キーを使用して設定をオーバーライドできます。Trino 内部認証の詳細については、[Trino 353 ドキュメントの「Secure internal communication」](#)を参照してください。

さらに、Amazon EMR リリースバージョン 5.10.0 以降では、HTTPS を使用した Presto コーディネーターへのクライアント接続に [LDAP 認証](#) を設定できます。このセットアップでは、安全な LDAP (LDAPS) を使用します。LDAP サーバーで TLS を有効にする必要があります。また、Presto クラスターは転送時のデータの暗号化を有効にして、セキュリティ設定を使用する必要があります。その他の設定も必要です。設定オプションは、使用する Amazon EMR のリリースバージョンに応じて異なります。詳細については、「[Amazon EMR の Presto に LDAP 認証を使用する](#)」を参照してください。

Amazon EMR の Presto は、デフォルトでは内部 HTTPS にポート 8446 を使用します。内部通信に使用するポートは、Presto コーディネーターへのクライアント HTTPS アクセスに使用するものと同じポートである必要があります。`http-server.https.port` 設定分類の `presto-config` プロパティでは、ポートが指定されます。

## Amazon EMR の Presto に LDAP 認証を使用する

このセクションのステップに従って、LDAP を設定します。例および詳細へのリンクについては、各ステップを参照してください。

### LDAP 認証を設定するステップ

- [ステップ 1: LDAP サーバーに関する情報を収集し、サーバー証明書を Amazon S3 にコピーする](#)
- [ステップ 2: セキュリティ設定をセットアップする](#)
- [ステップ 3: LDAP 用の Presto プロパティで設定 JSON を作成する](#)
- [ステップ 4: LDAP サーバー証明書をコピーするスクリプトを作成し、Amazon S3 にアップロードする](#)
- [ステップ 5: クラスターを作成する](#)

## ステップ 1: LDAP サーバーに関する情報を収集し、サーバー証明書を Amazon S3 にコピーする

LDAP 認証を設定するには、LDAP サーバーから次のセクションの情報と項目が必要になります。

### LDAP サーバーの IP アドレスまたはホスト名

Amazon EMR マスターノードの Presto コーディネーターは、指定された IP アドレスまたはホスト名で LDAP サーバーに到達できる必要があります。デフォルトでは、Presto はポート 636 経由で LDAPS を使用して LDAP サーバーと通信します。LDAP 実装でカスタムポートが必要になる場合は、Amazon EMR 5.16.0 以降で `ldap.url` プロパティを使用するか、以前のバージョンで `authentication.ldap.url` を使用してレポートを指定できます。「636」の `presto-config` の設定分類の例に示すように、[ステップ 3: LDAP 用の Presto プロパティで設定 JSON を作成する](#) のカスタムポートを置き換えます。ファイアウォールやセキュリティグループにより、ポート 636 (またはカスタムポート) およびポート 8446 (またはカスタムポート) でインバウンドおよびアウトバウンドトラフィックが許可されることを確認します。これは内部クラスター通信に使用されます。

### LDAP サーバー証明書

証明書ファイルを Amazon S3 の安全な場所にアップロードする必要があります。詳細については、「Amazon Simple Storage Service ユーザーガイド」の「[S3 バケットにファイルとフォルダをアップロードする方法](#)」を参照してください。クラスターの起動時に、この証明書を Amazon S3 からクラスター内の各ノードにコピーするブートストラップアクションを作成します。「[ステップ 4: LDAP サーバー証明書をコピーするスクリプトを作成し、Amazon S3 にアップロードする](#)」で、証明書の例は `s3://MyBucket/ldap_server.crt` です。

### 匿名バインド用の LDAP サーバーの設定

PrestoDB で匿名バインドが無効になっている場合は、LDAP サーバーにバインドし、PrestoDB サーバーが接続を確立できるようにするアクセス権限があるアカウントのユーザー ID (UID) とパスワードが必要です。UID とパスワードは、`internal-communication.authentication.ldap.user` 分類設定の `internal-communication.authentication.ldap.password` および `presto-config` プロパティを使用して指定します。Amazon EMR 5.10.0 はこれらの設定をサポートしないため、このリリースバージョンを使用するときは、匿名バインドが LDAP サーバーでサポートされている必要があります。

Trino では匿名バインディングを設定する必要はありません。

## LDAP サーバーで匿名バインドのステータスを取得するには

- 次の例に示すように、Linux クライアントから `ldapwhoami` コマンドを使用します。

```
ldapwhoami -x -H ldaps://LDAPServerHostNameOrIPAddress
```

匿名バインドが許可されない場合、コマンドの出力は次のようになります。

```
ldap_bind: Inappropriate authentication (48)
additional info: anonymous bind disallowed
```

アカウントに、シンプルな認証を使用する LDAP サーバーへのアクセス権限があることを確認するには

- 次の例に示すように、Linux クライアントから `ldapwhoami` コマンドを使用します。この例では、架空のホスト名 `ip-xxx-xxx-xxx-xxx.ec2.internal` の EC2 インスタンスで実行されている Open LDAP サーバーに保存されている架空のユーザー `presto` を使用しています。このユーザーは組織単位 (OU) `admins` とパスワード `123456` に関連付けられています。

```
ldapwhoami -x -w "123456" -D uid=presto,ou=admins,dc=ec2,dc=internal -H ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal
```

アカウントが有効であり、適切なアクセス権限がある場合、コマンドの出力は次のようになります。

```
dn:uid=presto,ou=admins,dc=ec2,dc=internal
```

[ステップ 3: LDAP 用の Presto プロパティで設定 JSON を作成する](#) の設定例には、わかりやすくするためにこのアカウントが含まれています。ただし、5.10.0 の例は例外で、サポートされていません。LDAP サーバーが匿名バインドを使用する場合は、`internal-communication.authentication.ldap.user` と `internal-communication.authentication.ldap.password` の名前と値のペアを削除します。

### Presto ユーザーの LDAP 識別名 (DN)

Presto の LDAP 設定を指定するときは、`${USER}` と組織単位 (OU)、および追加のドメインコンポーネント (DC) で構成されるバインドパターンを指定します。Presto は、パス



ワードの認証中に `${USER}` を各ユーザーの実際のユーザー ID (UID) に置き換え、このバインドパターンが指定する識別名 (DN) に一致させます。対象のユーザーが属している OU とその DC が必要です。たとえば、admins ドメインの corp.example.com OU のユーザーが Presto に対して認証されるようにするには、ユーザーバインドパターンとして `${USER},ou=admins,dc=corp,dc=example,dc=com` を指定します。

### Note

を使用する場合 AWS CloudFormation、を実際のユーザー ID (UID) `${USER}` に置き換えるには、`Fn::Sub` 関数を使用する必要があります。詳細については、「AWS CloudFormation ユーザーガイド」の「[Fn::Sub](#)」トピックを参照してください。

Amazon EMR 5.10.0 を使用する場合、このようなパターンを 1 つだけ指定できます。Amazon EMR 5.11.0 以降を使用して、複数のパターンをコロン (:) で区切って指定できます。Presto に対して認証を試みるユーザーは最初のパターンと比較されてから、2 番目のパターンに比較され、以降も同様に比較されます。例については、[ステップ 3: LDAP 用の Presto プロパティで設定 JSON を作成する](#)を参照してください。

## ステップ 2: セキュリティ設定をセットアップする

転送時の暗号化を有効にして、セキュリティ設定を作成します。詳細については、「Amazon EMR 管理ガイド」の「[セキュリティ設定を作成する](#)」を参照してください。転送時の暗号化を設定する際に指定する暗号化アーティファクトは、Presto のノード間の内部通信を暗号化するために使用されます。詳細については、「[伝送中のデータの暗号化のための証明書を提供する](#)」を参照してください。LDAP サーバー証明書は、Presto サーバーへのクライアント接続を認証するために使用されます。

## ステップ 3: LDAP 用の Presto プロパティで設定 JSON を作成する

LDAP 用の Presto プロパティを設定するには、`presto-config` 設定分類を使用します。`presto-config` の形式と内容は、Amazon EMR のリリースバージョンと Presto のインストール済み環境 (PrestoDB または Trino) によって若干異なります。設定の違いの例については、このセクションの後半で示します。詳細については、「[アプリケーションの設定](#)」を参照してください。

次の手順では、JSON データをファイル `MyPrestoConfig.json` に保存することを前提としています。コンソールを使用する場合、Amazon S3 の安全な場所にファイルをアップロードし、クラスターを作成するときに参照できるようにします。を使用する場合は AWS CLI、ファイルをローカルで参照できます。

## Example Amazon EMR 6.1.0 以降で PrestoSQL (Trino) を使用

次の例では、[ステップ 1: LDAP サーバーに関する情報を収集し、サーバー証明書を Amazon S3 にコピーする](#) の LDAP ホスト名を使用して、バインドのために LDAP サーバーに対して認証を行います。2つのユーザーバインドパターンが指定され、それにより、admins OU 内のユーザーと LDAP サーバーの datascientists OU は、ユーザーとして Trino サーバーへの認証の対象となります。バインドパターンはコロン (:) で区切られます。

Amazon EMR バージョン 6.4.0 以降では、PrestoSQL ではなく、新しい名称の Trino を使用します。Trino を使用する場合、以下の設定分類で *prestoql-config* を trino-config に置き換え、*prestoql-password-authenticator* を trino-password-authenticator に置き換えてください。

```
[
  {
    "Classification": "prestoql-config",
    "Properties": {
      "http-server.authentication.type": "PASSWORD"
    }
  },
  {
    "Classification": "prestoql-password-authenticator",
    "Properties": {
      "password-authenticator.name": "ldap",
      "ldap.url": "ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
      "ldap.user-bind-pattern": "uid=${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal"
    }
  }
]
```

## Example Amazon EMR 5.16.0 以降

次の例では、LDAP ユーザー ID とパスワード、および「[ステップ 1: LDAP サーバーに関する情報を収集し、サーバー証明書を Amazon S3 にコピーする](#)」の LDAP ホスト名を使用して、バインドのために LDAP サーバーに対して認証を行います。2つのユーザーバインドパターンが指定され、それにより、admins OU 内のユーザーと LDAP サーバーの datascientists OU は、ユーザーとして Presto サーバーへの認証の対象となります。バインドパターンはコロン (:) で区切られます。

```
[{
  "Classification": "presto-config",
```



```

        "Properties": {
            "http-server.authentication.type": "PASSWORD"
        }
    },
    {
        "Classification": "presto-password-authenticator",
        "Properties": {
            "password-authenticator.name": "ldap",
            "ldap.url": "ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
            "ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
            "internal-communication.authentication.ldap.user": "presto",
            "internal-communication.authentication.ldap.password": "123456"
        }
    }
}]]

```

### Example Amazon EMR 5.11.0 ~ 5.15.0

presto-config 設定分類の形式は、これらのリリースバージョンごとに少し異なります。次の例では、前の例と同じパラメータを指定します。

```

[ {
    "Classification": "presto-config",
    "Properties": {
        "http-server.authentication.type": "LDAP",
        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
        "authentication.ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
        "internal-communication.authentication.ldap.user": "presto",
        "internal-communication.authentication.ldap.password": "123456"
    }
} ]

```

### Example Amazon EMR 5.10.0

Amazon EMR 5.10.0 は匿名バインドのみをサポートするため、これらのエントリは省略されます。さらに、指定できるバインドパターンは 1 つのみです。

```

[ {
    "Classification": "presto-config",
    "Properties": {

```

```
        "http-server.authentication.type": "LDAP",
        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
        "ldap.user-bind-pattern": "uid=
${USER},ou=prestousers,dc=ec2,dc=internal"
    }
}]
```

## ステップ 4: LDAP サーバー証明書をコピーするスクリプトを作成し、Amazon S3 にアップロードする

クラスターの各ノードに証明書ファイルをコピーするスクリプトを作成し、キーストアに追加します。テキストエディタを使用してスクリプトを作成し、保存した後、Amazon S3 にアップロードします。では[ステップ 5: クラスターを作成する](#)、スクリプトファイルは `s3://MyBucket/LoadLDAPCert.sh` として参照されます。

次の例では、スクリプトはデフォルトのキーストアパスワード `changeit` を使用します。クラスターの作成後にマスターノードに接続し、`keytool` コマンドを使用してキーストアパスワードを変更することをお勧めします。

```
#!/bin/bash
aws s3 cp s3://MyBucket/ldap_server.crt .
sudo keytool -import -keystore /usr/lib/jvm/jre-1.8.0-openjdk.x86_64/lib/security/cacerts -trustcacerts -alias ldap_server -file ./ldap_server.crt -storepass changeit -noprompt
```

## ステップ 5: クラスターを作成する

クラスターを作成するときは、Presto と、Amazon EMR でインストールする他のアプリケーションを選択します。次の例では、JSON 内で設定分類プロパティを参照していますが、インラインで設定分類を指定することもできます。

Amazon EMR コンソールを使用して LDAP 認証により Presto クラスターを作成するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。

3. Presto と、Amazon EMR でインストールする他のアプリケーションを選択し、[ソフトウェア設定] で、使用する Amazon EMR の [リリース] を選択します。LDAP 認証は Amazon EMR 5.10.0 以降でのみサポートされます。
4. [Edit software settings] (ソフトウェア設定の編集) で [Load JSON from S3] (S3 からの JSON のロード) を選択し、「[ステップ 3: LDAP 用の Presto プロパティで設定 JSON を作成する](#)」で作成した JSON 設定ファイルの Amazon S3 の場所を入力し、[Next] (次へ) を選択します。
5. クラスターハードウェアおよびネットワークを設定し、[次へ] を選択します。
6. [ブートストラップアクション] を選択します。[ブートストラップアクションを追加する] で [カスタムアクション]、[設定と追加] の順に選択します。
7. ブートストラップアクションの名前を入力し、「」で作成したスクリプトの場所を入力します。例えば[ステップ 4: LDAP サーバー証明書をコピーするスクリプトを作成し、Amazon S3 にアップロードする](#)、s3:///LoadLDAPCert.sh と入力し、「追加」を選択します。  
MyBucketLoadLDAPCert.sh
8. [汎用オプション]、[タグ]、[追加のオプション] で、アプリケーションに最適な設定を選択し、[次へ] を選択します。
9. [認証と暗号化] を選択し、「」で作成した [セキュリティ設定[ステップ 2: セキュリティ設定をセットアップする](#)] を選択します。
10. アプリケーションに最適な他のセキュリティオプションを選択し、[クラスターの作成] を選択します。

AWS CLIを使用して LDAP 認証により Presto クラスターを作成するには

- `aws emr create-cluster` コマンドを実行します。少なくとも、Presto アプリケーションと、前のステップで作成した Presto 設定分類、ブートストラップスクリプト、およびセキュリティ設定を指定します。次の例では、コマンドを実行するのと同じディレクトリに JSON ファイルとして保存されている設定ファイルを参照します。一方、ブートストラップスクリプトは、Amazon S3 に保存する必要があります。次の例では `s3:///MyBucket/LoadLDAPCert.sh` を使用しています。

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --applications Name=presto --release-label emr-5.16.0 \
--use-default-roles --ec2-attributes KeyName=MyKeyPair,SubnetId=subnet-1234ab5
\ --instance-count 3 --instance-type m5.xlarge --region us-west-2 --name
"MyPrestoWithLDAPAuth" \
--bootstrap-actions Name="Distribute LDAP server cert",Path="s3://MyBucket/
LoadLDAPCert.sh" \
--security-configuration MyPrestoLDAPSecCfg --configurations file://
MyPrestoConfig.json
```

## Presto 厳格モードを有効にする

特定の状況では、クエリを長時間実行するとコストが高くなり、Amazon EMR がより多くのクラスターリソースを使用する可能性があります。これにより、クラスター上の他のワークロードからリソースが奪われます。Amazon EMR バージョン 6.8 以降では、次のような長時間実行されるクエリを拒否または警告する厳格モード機能を使用できます。

- パーティション化された列に述語がないクエリ。これにより大量のデータがテーブルスキャンされます
- 2つの大きなテーブルを相互結合するクエリ
- 多数の行を無制限にソートするクエリ

Presto がクエリプランを完全に最適化すると、厳格モードが実行されます。厳格モードを使用してクエリのニーズに合わせてカスタマイズするには、以下の方法で Presto を設定できます。

### 厳格モードのための Presto 設定

設定	説明	デフォルト値
strict-mode-enabled	厳格モードのオンとオフを切り替えます。true の値は厳格モードがオンであることを示します。	false
strict-mode-fail-query	厳格モードでは実行時間が長くなりそうなクエリが検出された場合、クエリを拒否しま	false

設定	説明	デフォルト値
<code>strict-mode-restrictions</code>	<p>す。false の場合、Amazon EMR は警告を発するだけです。</p> <p>厳格モードを有効にするときに適用する制限を指定します。厳格モードでは次の制限がサポートされます: MANDATORY_PARTITION_PREDICATE、DISALLOW_CROSS_JOIN、LIMITED_SORT。</p>	MANDATORY_PARTITION_PREDICATE、DISALLOW_CROSS_JOIN、LIMITED_SORT

厳格モードを試してみるには、これらの設定をオーバーライドするか、Presto クライアントを使用するときにセッションプロパティとして設定します。

を使用してクラスターの作成時に設定を設定するには AWS Management Console

1. [クラスターの作成] を選択し、Amazon EMR バージョン 6.8.0、および Presto または Trino を選択します。詳細については、「[PrestoDB と Trino をインストールする](#)」を参照してください。
2. 厳格モードの設定プロパティを直接指定するか、JSON ファイルを Amazon S3 にアップロードします。オプションで、メタストア用を選択します。VPC、サブネット、ブートストラップアクション、キーペア、セキュリティグループを指定します。クラスターを作成するには、[クラスターの作成] を選択します。
3. クラスターのプライマリノードにログインし、`presto-cli` または `trino-cli` を実行します。
4. クエリを送信します。厳格モードは各クエリを検証し、実行時間が長いかどうかを判断します。`strict-mode-fail-query` 設定に応じて、Amazon EMR はクエリを拒否するか、警告を發します。
5. クエリが完了したら、クラスターを終了し、リソースを削除します。

を使用して実行中のクラスターで 設定を設定するには AWS CLI

1. を使用してクラスターのプライマリノードにログイン AWS CLI し、 `presto-cli` または `trino-cli` を実行します。
2. 以下のコマンドを希望する値で実行します。

```
set session strict_mode_enabled = true;
set session strict_mode_fail_query = false;
set session strict_mode_restrictions = 'DISALLOW_CROSS_JOIN,LIMITED_SORT';
```

## 考慮事項

厳格モードを使用する場合は、次の点を考慮してください。

- 厳格モードでは、リソースをあまり消費しない短時間のクエリを拒否できる場合があります。例えば、小さなテーブルのクエリでは、動的フィルタリングは適用されず、内部結合をクロス結合に置き換えることもありません。これにより、クエリで必須パーティション述語が使用されたり、クロス結合が禁止されたりする可能性があります。この場合、厳格モードではクエリが拒否されます。
- 厳格モードのチェックは、SELECT、INSERT、CREATE TABLE AS SELECT、および EXPLAIN ANALYZE クエリタイプにのみ適用されます。
- 必須パーティション述語の制限は Hive コネクタでのみ使用できます。

## Presto でのスポットインスタンス損失の処理

Amazon EMR のスポットインスタンスを使用すると、予備の Amazon EC2 キャパシティでビッグデータワークロードを低コストで実行できます。Amazon EC2 は、コスト削減と引き換えに、2 分間の通知により、スポットインスタンスを中断できます。ノードを終了すると、Presto がエラーを返すまでに最大 10 分かかることがあります。これにより、エラー報告が不必要に遅れたり、再試行されたりする可能性があります。クイック終了は、終了したノードを Presto が処理する方法を制御できる機能です。

Presto コーディネーターの仕事は、定期的にステータスをポーリングしてすべてのワーカーノードを追跡することです。クイック終了がない場合、コーディネーターは各ノードのステータス NodeManager について YARN を参照しません。これにより、クエリが失敗するまでの再試行ループが長くなる可能性があります。クイックターミネーションでは、Presto コーディネーターは、ポーリングがホストに到達するのに失敗すると NodeManager すぐに、 のノードステータスを調べま

す。がノードが非アクティブである NodeManager ことを示す場合、Presto はそれ以上の再試行を中止し、クエリに失敗してNODE\_DECOMMISSIONEDエラーを返します。

以下の設定パラメータセットにより、ノード終了時の Presto の動作を制御およびカスタマイズできます。

### ノード障害処理用の Presto 設定

設定	説明	デフォルト値
<code>query.remote-task.max-backoff-duration</code>	コーディネーターがワーカーノードからリモートタスクのステータスを取得しようとし続ける時間。	10 分
<code>query.remote-task.quick-terminate-node-failure</code>	<p>コーディネーターがノードに到達できない場合や、そのノードで実行されているワーカーに接続できない場合に、クイックノード障害を有効にします。<code>query.remote-task.terminate-on-connect-exception</code> の値によって、コーディネーターがノードに到達する必要があるのか、ワーカーに接続する必要があるのかが決まります。</p> <p>ノードはクエリに失敗し、Amazon EMR は利用可能なワーカーのリストからノードを削除します。この場合、ノードを使用して新しいクエリをスケジュールすることはできません。</p> <p>この値を <code>false</code> に設定すると、Presto コーディネー</p>	true

設定	説明	デフォルト値
	ターが再びノードにアクセスしようとする (query.remote-task.max-backoff-duration の場合)、以前の動作に戻ります。その後、ノードが使用不可とマークされ、ノード上で実行中のクエリが失敗します。	
query.remote-task.terminate-on-connection-exception	ホストには接続可能だが、コーディネーターがホストのワーカプロセスに接続できない場合に Amazon EMR がノードを処理する必要があるかどうかを指定します。この値を true に設定すると、ホストにアクセスできない場合にクイッククエリ障害が有効になります。	false

## Trino での耐障害性実行

耐障害性実行は、クラスターがクエリの失敗を軽減するために使用できる Trino のメカニズムです。これを実現するには、クエリやそのコンポーネントタスクが失敗したときに再試行します。耐障害性実行が有効になると、中間交換データがスプールされ、クエリの実行中にワーカーが停止したり、その他の障害が発生したりした場合に、別のワーカーがそのデータを再利用できます。

Trino での耐障害性実行について詳しくは、Trino ブログの「[Project Tardigrade delivers ETL at Trino speeds to early users](#)」を参照してください。

### 構成

耐障害性実行はデフォルトでは無効になっています。この機能を有効にするには、以下のよう  
に、trino-config 分類の retry-policy 設定プロパティを、必要な再試行ポリシーに基づいて、QUERY または TASK に設定します。



```
{"classification":  
  "trino-config",  
"properties":  
  {  
    "retry-policy":  
      "QUERY"  
  }  
}
```

**QUERY** 再試行ポリシーは、ワーカーノードでエラーが発生したときに自動的にクエリを再試行するよう Trino に指示します。Trino クラスターのワークロードの大半が多数の小さなクエリで構成されている場合は、QUERY 再試行ポリシーを使用することをお勧めします。

**TASK** 再試行ポリシーは、障害発生時に個々のクエリタスクを再試行するよう Trino に指示します。Trino が大量のバッチクエリを実行する場合にはこのポリシーをお勧めします。クラスターは、クエリ全体を再試行するよりも、クエリ内の小さなタスクの方がより効率的に再試行できます。

## 交換マネージャー

交換マネージャーは、耐障害性実行のためにスプールされたデータを保存および管理します。外部ストレージを使用して、メモリ内のバッファサイズを超えてオーバーフローしたデータを保存します。スプールされたデータを Amazon S3、Amazon S3 互換システム、HDFS などの指定された場所に保存するファイルシステムベースの交換マネージャーを設定できます。

Amazon EMR リリース 6.9.0 以降には、交換マネージャーを設定するための `trino-exchange-manager` 分類が含まれています。これらのリリースでは、HDFS によるスプールもサポートされています。

### 交換マネージャーの設定

`trino-exchange-manager` 設定分類を使用して交換マネージャーを設定します。この分類により、コーディネーターとすべてのワーカーノードに内部的に `etc/exchange-manager.properties` 設定ファイルが作成されます。またこの分類により、`exchange-manager.name` 設定プロパティは `filesystem` に設定されます。

デフォルトでは、Amazon EMR リリース 6.9.0 以降は HDFS を交換マネージャーとして使用します。HDFS は Amazon EMR EC2 クラスターで使用でき、スプールはデフォルトで `trino-exchange/` ディレクトリで行われます。デフォルトの設定を使用するには、次の設定を行います。

```
{"Classification":
```

```
"trino-exchange-manager"
}
```

カスタムの場所を指定する場合は、trino-exchange-manager 分類で次のプロパティを設定します。

- exchange.use-local-hdfs を true に設定します。
- exchange.base-directories を HDFS 内のカスタムディレクトリの場所 (例: exchange.base-directories=/exchange) に設定します。カスタムディレクトリがまだ HDFS にはない場合は、Amazon EMR によって作成されます。

## HDFS 交換マネージャーの設定

内部テストの結果によると、他のクラウドベースのファイルシステムと比較してクエリパフォーマンスを向上させるには、ローカル HDFS にスプールすることをお勧めします。HDFS を使用する交換マネージャーには、次の設定を行うことができます。

構成	説明	デフォルト設定
exchange.hdfs.block-size	HDFS ストレージのブロックサイズ	4 MB
hdfs.config.resources	HDFS を設定するファイルパスのリスト	exchange.use-local-hdfs が true の場合は core-site.xml 、hdfs-site.xml ファイルへのパスを使用し、それ以外の場合は null

その他の耐障害性実行設定プロパティ、および Amazon S3 またはその他の Amazon S3 互換システムをスプール用に設定する方法については、Trino ドキュメントの「[Fault-tolerant execution](#)」ページを参照してください。

## 考慮事項と制約事項

- 耐障害性実行を有効にすると、retry-policy が設定されているときに write をサポートしていないコネクタの write 操作は無効になります。Amazon EMR リリース 6.9.0 では、Delta

Lake、Hive および Iceberg の各コネクタは `retry-policy` による `write` 操作をサポートしています。

- 交換マネージャーを使用して高コストの I/O 操作を実行すると、交換マネージャーが中間データを外部ストレージにスプールする間、クエリのパフォーマンスが低下する可能性があります。

## グレースフルな廃止による Presto Auto Scaling の使用

Amazon EMR リリースバージョン 5.30.0 以降には、特定のスケールアップアクションの猶予期間を設定する機能が含まれています。猶予期間を使用することで、スケールインサイズ変更アクションまたは自動スケールアップポリシーリクエストによってノードが終了するまで Presto タスクの実行を続けることができます。スケールアップルールの詳細については、「Amazon EMR 管理ガイド」の「[自動スケールアップルールについて](#)」を参照してください。グレースフルな廃止を使用する Presto Auto Scaling では、廃止処理中のノードでの新しいタスクのスケジューリングを回避するとともに、シャットダウンタイムアウトに達する前に、既に実行中のタスクを完了できます。実行中のクエリは、ノードが廃止される前に実行を完了します。Autoscaling は、インスタンスフリートではサポートされていません。

自動スケールアップのシャットダウンリクエストが受け取られた後、Presto タスクが完了するまでの許容時間を制御できます。デフォルトでは、Amazon EMR のシャットダウンタイムアウトは 0 分です。この場合、スケールインリクエストで必要になると、Amazon EMR によってノードとそのノードで実行されている Presto タスクが直ちに終了します。実行中のクエリをクラスターのスケールダウン前に完了できるように、Amazon EMR で Presto タスクのタイムアウトを長く設定するには、`presto-config` 設定分類を使用して、`graceful-shutdown-timeout` パラメータを 0 より大きい値 (秒単位または分単位) に設定します。詳細については、「[アプリケーションの設定](#)」を参照してください。

例えば、`graceful-shutdown-timeout` 値を "30m" に増加すると、30 分間のタイムアウト期間が指定されます。シャットダウンタイムアウト期間が終了すると、廃止対象としてマークされたノードは、クエリタスクの完了を待機していても強制的に終了され、クエリは失敗します。クエリタスクが 5 分後に終了した場合、他の YARN アプリケーションが実行を完了していれば、廃止対象としてマークされたノードは終了します。

### Example グレースフルな廃止による Presto 自動スケールアップ設定の例

`graceful-shutdown-timeout` 値を、設定に適した分数に置き換えます。決められた最大値はありません。次の例では、タイムアウト値を 1800 秒 (30 分) に設定します。

[

```
{
  "classification": "presto-config",
  "properties": {
    "graceful-shutdown-timeout": "1800s"
  }
}
```

## 制約事項

PrestoDB のグレースフルな廃止は、HTTP 接続が無効になっている EMR クラスターでは動作しません (`http-server.http.enabled` が `false` に設定されている場合など)。Trino では、`http-server.http.enabled` の設定に関係なく、グレースフルな廃止はサポートされていません。

## Amazon EMR での Presto に関する考慮事項

Amazon EMR での [Presto](#) を実行するときは、次の制限について考慮してください。

### Presto コマンドライン実行可能ファイル

Amazon EMR では、以下の例に示されているように、PrestoDB と Trino は両方とも同じコマンドライン実行可能ファイル `presto-cli` を使用します。

```
presto-cli --catalog hive
```

### Presto デプロイプロパティは設定可能ではない

ご使用の Amazon EMR のバージョンによって、使用可能な Presto デプロイ設定が決まります。これらの設定プロパティについては、Presto ドキュメントの「[Deploying Presto](#)」を参照してください。次の表に、Presto properties ファイルのさまざまな設定オプションを示します。

File	設定可能
<code>log.properties</code>	PrestoDB: Amazon EMR バージョン 4.0.0 以降で設定可能です。presto-log 設定分類を使用します。

File	設定可能
	Trino (PrestoSQL): Amazon EMR バージョン 6.1.0 以降で設定可能です。prestoql-log または trino-log 設定分類を使用します。
config.properties	<p>PrestoDB: Amazon EMR バージョン 4.0.0 以降で設定可能です。presto-config 設定分類を使用します。</p> <p>Trino (PrestoSQL): Amazon EMR バージョン 6.1.0 以降で設定可能です。prestoql-config または trino-config 設定分類を使用します。</p>
hive.properties	<p>PrestoDB: Amazon EMR バージョン 4.1.0 以降で設定可能です。presto-connector-hive 設定分類を使用します。</p> <p>Trino (PrestoSQL): Amazon EMR バージョン 6.1.0 以降で設定可能です。prestoql-connector-hive または trino-connector-hive 設定分類を使用します。</p>
node.properties	<p>PrestoDB: Amazon EMR バージョン 5.6.0 以降で設定可能です。presto-node 設定分類を使用します。</p> <p>Trino (PrestoSQL): Amazon EMR バージョン 6.1.0 以降で設定可能です。prestoql-node または trino-node 設定分類を使用します。</p>
jvm.config	設定可能ではありません。

## PrestoDB と Trino のインストール

アプリケーション名 Presto は、クラスターに PrestoDB をインストールするために引き続き使用されます。クラスターに Trino をインストールするには、アプリケーション名 Trino (または Amazon EMR の古いバージョンでは PrestoSQL) を使用します。

PrestoDB または Trino のいずれかをインストールできますが、両方を単一のクラスターにインストールすることはできません。クラスターを作成しようとしたときに PrestoDB と Trino の両方を指定すると、検証エラーが発生し、クラスター作成リクエストは失敗します。

## EMRFS と PrestoS3FileSystem の設定

Amazon EMR バージョン 5.12.0 以降では、PrestoDB は EMRFS を使用できます。これはデフォルトの設定です。EMRFS は、Amazon EMR バージョン 6.1.0 以降では、デフォルトのファイルシステムまたは Trino (PrestoSQL) でもあります。詳細については、「Amazon EMR 管理ガイド」の「[EMR ファイルシステム \(EMRFS\)](#)」を参照してください。以前のバージョンの Amazon EMR では、PrestoS3 FileSystem が唯一の設定オプションです。

セキュリティ設定を使用して、Amazon S3 で EMRFS データの暗号化を設定できます。また、Amazon S3 への EMRFS リクエストに IAM ロールを使用することもできます。詳細については、「Amazon EMR 管理ガイド」の「[暗号化オプションを理解する](#)」および「[Amazon S3 への EMRFS リクエストの IAM ロールを設定する](#)」を参照してください。

### Note

Amazon EMR バージョン 5.12.0 を使用して Amazon S3 で基盤となるデータをクエリするときに、Presto エラーが発生することがあります。これは、Presto が `emrfs-site.xml` から設定分類値を取得できないためです。回避策として、`usr/lib/presto/plugin/hive-hadoop2/` の下に `emrfs` サブディレクトリを作成し、既存の `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` ファイルへのシンボリックリンクを `usr/lib/presto/plugin/hive-hadoop2/emrfs` に作成します。次に `presto-server` プロセスを再起動します (`sudo presto-server stop` に続いて `sudo presto-server start`)。

EMRFS のデフォルトを上書きし、代わりに PrestoS3FileSystem を使用できます。これを行うには、次の例に示すように `presto-connector-hive` 設定分類を使用して、`hive.s3-file-system-type` を `PRESTO` に設定します。詳細については、「[アプリケーションの設定](#)」を参照してください。

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.s3-file-system-type": "PRESTO"
    }
  }
]
```

PrestoS3FileSystem を使用する場合は、presto-connector-hive 設定分類 または trino-connector-hive Trino を使用して PrestoS3 FileSystem プロパティを設定します。使用可能なプロパティの詳細については、Presto ドキュメントの「Hive Connector」セクションにある「[Amazon S3 の設定](#)」を参照してください。これらの設定は EMRFS には適用されません。

## エンドユーザー偽装のデフォルト設定

デフォルトでは、Amazon EMR バージョン 5.12.0 以降では、エンドユーザーが HDFS にアクセスするための偽装を行うことができます。詳細については、Presto のドキュメントの「[エンドユーザー偽装](#)」を参照してください。この設定を presto-config 設定分類を使用して変更するには、hive.hdfs.impersonation.enabled プロパティを false に設定します。

## Presto ウェブインターフェイス用のデフォルトポート

デフォルトでは、Amazon EMR はポート 8889 を使用するように Presto コーディネーターの Presto ウェブインターフェイスを設定します (PrestoDB および Trino の場合)。このポートを変更するには、presto-config 設定分類を使用して http-server.http.port プロパティを設定します。詳細については、Presto ドキュメントの「Deploying Presto」セクションの「[Config properties](#)」を参照してください。

## 一部のリリースでの Hive バケットの実行に関する問題

Presto バージョン 152.3 には、特定の状況で Presto のクエリパフォーマンスを大幅に低下させる Hive バケットの実行に問題があります。Amazon EMR リリース 5.0.3、5.1.0、および 5.2.0 には、このバージョンの Presto が含まれています。この問題を軽減するには、次の例に示すように presto-connector-hive 設定分類を使用して、hive.bucket-execution プロパティを false に設定します。

```
[
```

```

{
  "Classification": "presto-connector-hive",
  "Properties": {
    "hive.bucket-execution": "false"
  }
}
]

```

## Presto リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Presto のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Presto バージョン情報

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.36.2	0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.1.0	0.284	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-



Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
		nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.0.0	0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.15.0	0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.14.0	0.281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.13.0	0.281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.12.0	0.281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.11.1	0.279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.11.0	0.279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.10.1	0.278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.10.0	0.278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.9.1	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.9.0	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.8.1	0.273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.8.0	0.273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.7.0	0.272	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.36.1	0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.36.0	0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker



Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.6.0	0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.35.0	0.266	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.5.0	0.261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.4.0	0.254.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.3.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.3.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.2.1	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.2.0	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.1.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.1.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-6.0.1	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.0.0	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.34.0	0.261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.33.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.33.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.32.1	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker



Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.32.0	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.31.1	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.31.0	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.30.2	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.30.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.30.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.29.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.28.1	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.28.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.27.1	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.27.0	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.26.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.25.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.24.1	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.24.0	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.23.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker



Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.23.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.22.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.21.2	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.21.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.21.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.20.1	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.20.0	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.19.1	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.19.0	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.18.1	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.18.0	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.17.2	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.17.1	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.17.0	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.16.1	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.16.0	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker



Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.15.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.15.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.14.2	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.14.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.14.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.13.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.13.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.3	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.12.2	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.1	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.12.0	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.4	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.11.3	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.2	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.11.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker



Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.10.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.10.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.9.1	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.9.0	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.8.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.8.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.7.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.7.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.6.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.6.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.4	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.5.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.5.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.4.1	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.4.0	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker



Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.3.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.3.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.3.0	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.2.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.2.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.1.1	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.1.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.0.3	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.2	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.1	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR リリースラベル	Presto バージョン	Presto でインストールされるコンポーネント
emr-5.0.0	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

次の表は、Amazon EMR の各リリースバージョンに含まれている Trino (Presto SQL) のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。バージョン 351以降、PrestoSQL は Trino に名称が変更されました。

#### Trino (PrestoSQL) バージョン情報

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-7.1.0	435	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-7.0.0	426	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode,

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
		hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.15.0	426	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.14.0	422	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.13.0	414	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.12.0	414	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.11.1	410	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.11.0	410	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker



Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.10.1	403	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.10.0	403	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.9.1	398	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.9.0	398	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.8.1	388	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.8.0	388	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.7.0	378	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.6.0	367	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.5.0	360	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker
emr-6.4.0	359	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.3.1	350	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-prestosql, hcatalog-server, mariadb-server, prestosql-coordinator, prestosql-worker
emr-6.3.0	350	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-prestosql, hcatalog-server, mariadb-server, prestosql-coordinator, prestosql-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.2.1	343	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-prestosql, hcatalog-server, mariadb-server, prestosql-coordinator, prestosql-worker
emr-6.2.0	343	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-prestosql, hcatalog-server, mariadb-server, prestosql-coordinator, prestosql-worker

Amazon EMR リリースラベル	Trino (PrestoSQL) バージョン	Trino (PrestoSQL) でインストールされるコンポーネント
emr-6.1.1	338	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-prestosql, hcatalog-server, mariadb-server, prestosql-coordinator, prestosql-worker
emr-6.1.0	338	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-prestosql, hcatalog-server, mariadb-server, prestosql-coordinator, prestosql-worker

## Trino (PrestoSQL) リリースノート (バージョン別)

- [Amazon EMR 6.9.0 - Trino \(PrestoSQL\) リリースノート](#)



## Amazon EMR 6.9.0 - Trino (PrestoSQL) リリースノート

### Amazon EMR 6.9.0 - Trino (PrestoSQL) の新機能

- 長時間実行されるクエリをサポートするため、Trino には耐障害性実行メカニズムが組み込まれました。耐障害性実行では、失敗したクエリやそのコンポーネントタスクを再試行することで、クエリの失敗を軽減します。詳細については、「[Trino での耐障害性実行](#)」を参照してください。

### Amazon EMR 6.9.0 - Trino (PrestoSQL) の変更点

#### Amazon EMR 6.9.0 - PrestoDB の変更

型	説明
Upgrade	PrestoDB の 0.276 へのアップグレード
Upgrade	Hadoop 3.3.3 をサポート
Upgrade	Hudi の 0.12.1 へのアップグレード
機能	GCSC API を使用したインタラクティブなワークロード AWS Lake Formation のための Amazon EMR と Presto の統合。
機能	Keberos を有効にするために、PrestoDB のセキュリティ設定内に Kerberos 関連の設定が追加されました。
バグ修正	hdfsConfiguration のコピー数を減らすために追加された OSS プルリクエスト #18115 を元に戻しました。これにより、EMRFS または Hudi テーブルを使用すると、HDFS 設定が正しくコピーされませんでした。

#### Amazon EMR 6.9.0 - Trino の変更

型	説明
Upgrade	Trino の 398 へのアップグレード

型	説明
Upgrade	Hadoop 3.3.3 をサポート
機能	Tardigrade のサポート: HDFS と Amazon S3 での交換スプーリングのサポートを追加しました。詳細については、「 <a href="#">Trino での耐障害性実行</a> 」を参照してください。
バグ修正	Trino Iceberg を使用していて、Glue カタログが有効になっている場合は、メタストア URI を <code>iceberg.properties.</code> に追加しません

### Amazon EMR 6.9.0 - Trino (PrestoSQL) の既知の問題点

- Amazon EMR リリース 6.9.0 では、Apache Ranger が有効になっているクラスターでは Trino は動作しません。Ranger で Trino を使用する必要がある場合は、[AWS Support](#) にお問い合わせください。

# Apache Spark

[Apache Spark](#) は、Amazon EMR クラスターでの機械学習、ストリーム処理、グラフ分析に役立つ分散処理フレームワークおよびプログラミングモデルです。Apache Spark は、Apache Hadoop と同様に、ビッグデータのワークロードを処理するために一般的に使用されているオープンソースの分散処理システムです。ただし、Spark には Hadoop といくつかの大きな違いがあります。MapReduce。Spark は、最適化された Directed Acyclic Graph (DAG) 実行エンジンを備えており、データをインメモリにアクティブにキャッシュするため、特に特定のアルゴリズムやインタラクティブクエリの場合にパフォーマンスが向上します。

Spark アプリケーションは、Scala、Java、および Python をネイティブでサポートしています。また、SQL ([Spark SQL](#))、機械学習 ([MLlib](#))、ストリーム処理 ([Spark Streaming](#))、グラフ処理 ([GraphX](#)) 用の緊密に統合されたライブラリもいくつか含まれています。これらのツールを使用すると、さまざまなユースケースで Spark フレームワークを活用しやすくなります。

Spark は、他の Hadoop アプリケーションと同時に Amazon EMR クラスターにインストールすることができ、Amazon EMR ファイルシステム (EMRFS) を利用して Amazon S3 のデータに直接アクセスすることができます。Hive は Spark と統合されているため、HiveContext オブジェクトを使用して Spark を使用して Hive スクリプトを実行できます。Hive コンテキストは、spark-shell に sqlContext として含められます。

Spark で EMR クラスターをセットアップし、サンプルデータセットを分析するチュートリアルについては、AWS ニュースブログの「[チュートリアル: Amazon EMR の開始方法](#)」を参照してください。

## Important

Apache Spark バージョン 2.3.1 は Amazon EMR リリース 5.16.0 以降で利用でき、[CVE-2018-8024](#) および [CVE-2018-1334](#) に対処します。Spark の以前のバージョンを Spark バージョン 2.3.1 以降に移行することをお勧めします。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Spark のバージョンと、Amazon EMR で Spark と共にインストールされるコンポーネントを示しています。

このリリースで Spark と共にインストールされるコンポーネントのバージョンについては、[「リリース 7.1.0 コンポーネントバージョン」](#)を参照してください。

## emr-7.1.0 の Spark バージョン情報

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-7.1.0	Spark 3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Spark のバージョンと、Amazon EMR で Spark と共にインストールされるコンポーネントを示しています。

このリリースで Spark と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の Spark バージョン情報

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.15.0	Spark 3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
		kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

### Note

Amazon EMR リリース 6.8.0 には、Apache Spark 3.3.0 が付属しています。この Spark リリースでは、Apache Log4j 2 と log4j2.properties ファイルを使用して Spark プロセス内の Log4j を設定します。クラスターで Spark を使用するか、カスタム設定パラメータを使用して EMR クラスターを作成し、Amazon EMR リリース 6.8.0 にアップグレードする場合は、Apache Log4j 2 の新しい spark-log4j2 設定分類とキー形式に移行する必要があります。詳細については、「[Apache Log4j 1.x から Log4j 2.x への移行](#)」を参照してください。

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Spark のバージョンと、Amazon EMR で Spark と共にインストールされるコンポーネントを示しています。

このリリースで Spark と共にインストールされるコンポーネントのバージョンについては、「[リリース 5.36.2 コンポーネントバージョン](#)」を参照してください。

### emr-5.36.2 の Spark バージョン情報

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.36.2	Spark 2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
		hdfs-namenode, hadoop- httpfs-server, hadoop-kms- server, hadoop-yarn-nodema- nager, hadoop-yarn-resour- cemanager, hadoop-yarn- timeline-server, hudi, hudi- spark, livy-server, nginx, r, spark-client, spark-history-serv- er, spark-on-yarn, spark-yarn- slave

## トピック

- [Apache Spark を使用したクラスターの作成](#)
- [Amazon EMR 6.x を使用して Docker で Spark アプリケーションを実行する](#)
- [AWS Glue データカタログを Spark SQL のメタストアとして使用する](#)
- [Spark の設定](#)
- [Spark パフォーマンスの最適化](#)
- [Spark 結果フラグメントキャッシュ](#)
- [Apache Spark 用の Nvidia RAPIDS アクセラレータの使用](#)
- [Spark シェルにアクセスする](#)
- [機械学習に Amazon SageMaker Spark を使用する](#)
- [Spark アプリケーションを作成する](#)
- [Amazon S3 で Spark のパフォーマンスを向上させる](#)
- [Spark ステップを追加する](#)
- [Spark アプリケーション履歴を表示する](#)
- [Spark ウェブ UI にアクセスする](#)
- [Spark 構造化ストリーミング Amazon Kinesis Data Streams コネクタの使用](#)
- [Amazon EMR での Apache Spark 用の Amazon Redshift インテグレーションの使用](#)
- [Spark リリース履歴](#)

## Apache Spark を使用したクラスターの作成

次の手順では、Amazon EMR コンソールで [クイックオプション] を使用して、[Spark](#) がインストールされたクラスターを作成します。

[Advanced Options] (詳細オプション) を使用してクラスター設定を詳細にカスタマイズしたり、アプリケーションをプログラムでインストールするステップを送信してカスタムアプリケーションを実行したりすることもできます。どちらのクラスター作成オプションでも、Spark SQL メタストアとして AWS Glue を使用するように選択できます。詳細については、「[AWS Glue データカタログを Spark SQL のメタストアとして使用する](#)」を参照してください。

Spark がインストールされたクラスターを起動するには

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Create cluster] (クラスターの作成) を選択して、[Quick Options] (クイックオプション) を使用します。
3. [Cluster name] (クラスター名) を入力します。クラスター名に <、>、\$、|、` (バックティック) の文字を含めることはできません。
4. [Software Configuration] (ソフトウェア設定) では、[Release] (リリース) オプションを選択します。
5. [Applications] (アプリケーション) では、Spark アプリケーションバンドルを選択します。
6. 必要に応じて他のオプションを選択し、[Create cluster (クラスターの作成)] を選択します。

### Note

クラスターを作成する場合に Spark を設定するには、「[Spark の設定](#)」を参照してください。

を使用して Spark がインストールされたクラスターを起動するには AWS CLI

- 次のコマンドを使用してクラスターを作成します。

```
aws emr create-cluster --name "Spark cluster" --release-label emr-7.1.0 --
applications Name=Spark \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 --use-
default-roles
```

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

SDK for Java を使用し、Spark がインストールされたクラスターを起動するには

SupportedProductConfig で使用される RunJobFlowRequest を使用して Spark をアプリケーションとして指定します。

- 次の例は、Java を使用した Spark のあるクラスターを作成する方法を示します。

```
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;
import com.amazonaws.services.elasticmapreduce.util.StepFactory;

public class Main {

    public static void main(String[] args) {
        AWSCredentials credentials_profile = null;
        try {
            credentials_profile = new
ProfileCredentialsProvider("default").getCredentials();
        } catch (Exception e) {
            throw new AmazonClientException(
                "Cannot load credentials from .aws/
credentials file. " +
                "Make sure that the
credentials file exists and the profile name is specified within it.",
                e);
        }

        AmazonElasticMapReduce emr =
AmazonElasticMapReduceClientBuilder.standard()
```



```
        .withCredentials(new
AWSStaticCredentialsProvider(credentials_profile))
        .withRegion(Regions.US_WEST_1)
        .build();

// create a step to enable debugging in the AWS Management Console
StepFactory stepFactory = new StepFactory();
StepConfig enableddebugging = new StepConfig()
    .withName("Enable debugging")
    .withActionOnFailure("TERMINATE_JOB_FLOW");

.withHadoopJarStep(stepFactory.newEnableDebuggingStep());

Application spark = new Application().withName("Spark");

RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("Spark Cluster")
    .withReleaseLabel("emr-5.20.0")
    .withSteps(enableddebugging)
    .withApplications(spark)
    .withLogUri("s3://path/to/my/logs/")
    .withServiceRole("EMR_DefaultRole")
    .withJobFlowRole("EMR_EC2_DefaultRole")
    .withInstances(new JobFlowInstancesConfig()
        .withEc2SubnetId("subnet-12ab3c45")
        .withEc2KeyName("myEc2Key")
        .withInstanceCount(3)

.withKeepJobFlowAliveWhenNoSteps(true)
        .withMasterInstanceType("m4.large")

.withSlaveInstanceType("m4.large"));
RunJobFlowResult result = emr.runJobFlow(request);
System.out.println("The cluster ID is " + result.toString());
    }
}
```

# Amazon EMR 6.x を使用して Docker で Spark アプリケーションを実行する

Amazon EMR 6.0.0 では、Spark アプリケーションで Docker コンテナを使用してライブラリの依存関係を定義できます。クラスター内の各 Amazon EC2 インスタンスに依存関係をインストールする必要はありません。Docker で Spark を実行するには、まず Docker レジストリを設定し、Spark アプリケーションの送信時に追加のパラメータを定義する必要があります。詳細については、「[Docker 統合の設定](#)」を参照してください。

アプリケーションが送信されると、YARN は Docker を呼び出して指定された Docker イメージをプルし、Docker コンテナ内で Spark アプリケーションを実行します。これにより、依存関係を簡単に定義して分離できます。結果として、ジョブ実行に必要なライブラリを使用して Amazon EMR クラスター内のインスタンスをブートストラップまたは準備する時間が短縮されます。

## Docker で Spark を実行する際の考慮事項

Docker で Spark を実行するときは、次の前提条件を満たす必要があります。

- docker パッケージと CLI は、コアノードとタスクノードにのみインストールします。
- Amazon EMR 6.1.0 以降では、次のコマンドを使用してプライマリノードに Docker をインストールすることもできます。

```
sudo yum install -y docker
sudo systemctl start docker
```

- spark-submit コマンドは、常に Amazon EMR クラスターのプライマリインスタンスから実行します。
- Docker イメージの解決に使用する Docker レジストリは、分類 API を使用して定義する必要があります。クラスターの起動時に container-executor 分類キーを使用して追加のパラメータを定義します。
  - docker.trusted.registries
  - docker.privileged-containers.registries
- Docker コンテナで Spark アプリケーションを実行するには、次の設定オプションが必要です。
  - YARN\_CONTAINER\_RUNTIME\_TYPE=docker
  - YARN\_CONTAINER\_RUNTIME\_DOCKER\_IMAGE={DOCKER\_IMAGE\_NAME}
- Amazon ECR を使用して Docker イメージを取得する場合は、クラスター自体を認証するようにクラスターを設定する必要があります。そのためには、次の設定オプションを使用します。

- YARN\_CONTAINER\_RUNTIME\_DOCKER\_CLIENT\_CONFIG={DOCKER\_CLIENT\_CONFIG\_PATH\_ON
- Amazon EMR 6.1.0 以降では、ECR 自動認証機能が有効になっている場合、リストされているコマンド  
YARN\_CONTAINER\_RUNTIME\_DOCKER\_CLIENT\_CONFIG={DOCKER\_CLIENT\_CONFIG\_PATH\_ON\_HDFS} を使用する必要はありません。
- Spark で Docker イメージを使用する場合、Docker イメージに Java をインストールする必要があります。

前提条件の詳細については、「[Docker 統合の設定](#)」を参照してください。

## Docker イメージの作成

Docker イメージを作成するには Dockerfile を使用します。このファイルは、イメージに含めるパッケージと設定を定義します。次の 2 つの Dockerfiles の例では、PySpark と SparkR を使用しています。

### PySpark Dockerfile

この Dockerfile から作成された Docker イメージには、Python 3 と NumPy Python パッケージが含まれています。この Dockerfile では、Amazon Linux 2 と Amazon Corretto JDK 8 を使用します。

```
FROM amazoncorretto:8

RUN yum -y update
RUN yum -y install yum-utils
RUN yum -y groupinstall development

RUN yum list python3*
RUN yum -y install python3 python3-dev python3-pip python3-virtualenv

RUN python -V
RUN python3 -V

ENV PYSPARK_DRIVER_PYTHON python3
ENV PYSPARK_PYTHON python3

RUN pip3 install --upgrade pip
RUN pip3 install numpy pandas

RUN python3 -c "import numpy as np"
```

## SparkR Dockerfile

この Dockerfile から作成した Docker イメージには、R と randomForest CRAN パッケージが含まれます。この Dockerfile には、Amazon Linux 2 と Amazon Corretto JDK 8 が含まれます。

```
FROM amazoncorretto:8

RUN java -version

RUN yum -y update
RUN amazon-linux-extras install R4

RUN yum -y install curl hostname

#setup R configs
RUN echo "r <- getOption('repos'); r['CRAN'] <- 'http://cran.us.r-project.org';
  options(repos = r);" > ~/.Rprofile

RUN Rscript -e "install.packages('randomForest')"
```

Dockerfile の構文の詳細については、[Dockerfile リファレンスのドキュメント](#)を参照してください。

## Amazon ECR の Docker イメージを使用する

Amazon Elastic Container Registry (Amazon ECR) は、フルマネージド型の Docker コンテナレジストリであり、Docker コンテナイメージの保存、管理、デプロイを容易に行うことができます。Amazon ECR を使用する場合は、ECR のインスタンスを信頼するようにクラスターを設定する必要があります。また、クラスターが Amazon ECR の Docker イメージを使用するように認証を設定する必要があります。詳細については、「[Amazon ECR にアクセスするための YARN の設定](#)」を参照してください。

Amazon ECR に保存されているイメージに Amazon EMR ホストがアクセスできるようにするには、インスタンスプロファイルに関連付けられた AmazonEC2ContainerRegistryReadOnly ポリシーからのアクセス許可がクラスターに必要です。詳細については、「[AmazonEC2ContainerRegistryReadOnly ポリシー](#)」を参照してください。

この例では、Amazon ECR レジストリが信頼されるように、次の追加設定を使用してクラスターを作成する必要があります。*123456789123.dkr.ecr.us-east-1.amazonaws.com* エンドポイントを、実際の Amazon ECR エンドポイントに置き換えます。

[

```
{
  "Classification": "container-executor",
  "Configurations": [
    {
      "Classification": "docker",
      "Properties": {
        "docker.privileged-containers.registries":
"local,centos,123456789123.dkr.ecr.us-east-1.amazonaws.com",
        "docker.trusted.registries": "local,centos,123456789123.dkr.ecr.us-
east-1.amazonaws.com"
      }
    }
  ],
  "Properties": {}
}
```

## Amazon ECR PySpark での の使用

次の例では、Amazon ECR にタグ付けおよびアップロードされる PySpark Dockerfile を使用しています。Dockerfile をアップロードしたら、PySpark ジョブを実行し、Amazon ECR から Docker イメージを参照できます。

クラスターを起動したら、SSH を使用してコアノードに接続し、次のコマンドを実行して Dockerfile の例からローカル PySpark Docker イメージを構築します。

まず、ディレクトリと Dockerfile を作成します。

```
mkdir pyspark
vi pyspark/Dockerfile
```

PySpark Dockerfile の内容を貼り付け、次のコマンドを実行して Docker イメージを構築します。

```
sudo docker build -t local/pyspark-example pyspark/
```

例の `emr-docker-examples` ECR リポジトリを作成します。

```
aws ecr create-repository --repository-name emr-docker-examples
```

ローカルに構築したイメージにタグを付けて ECR にアップロードし、[123456789123.dkr.ecr.us-east-1.amazonaws.com](https://123456789123.dkr.ecr.us-east-1.amazonaws.com) を実際の ECR エンドポイントに置き換えます。

```
sudo docker tag local/pyspark-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:pyspark-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:pyspark-example
```

SSH を使用してプライマリノードに接続し、main.py というファイル名で Python スクリプトを準備します。次の内容を main.py ファイルに貼り付け、保存します。

```
from pyspark.sql import SparkSession
spark = SparkSession.builder.appName("docker-numpy").getOrCreate()
sc = spark.sparkContext

import numpy as np
a = np.arange(15).reshape(3, 5)
print(a)
```

Amazon EMR 6.0.0 でジョブを送信するには、Docker イメージの名前を参照します。追加の設定パラメータを定義して、ジョブの実行で必ず Docker ガランタイムとして使用されるようにします。Amazon ECR を使用する場合、YARN\_CONTAINER\_RUNTIME\_DOCKER\_CLIENT\_CONFIG は Amazon ECR への認証に使用する認証情報が含まれている config.json ファイルを参照する必要があります。

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:pyspark-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=$DOCKER_CLIENT_CONFIG \
--num-executors 2 \
main.py -v
```

Amazon EMR 6.1.0 以降でジョブを送信するには、Docker イメージの名前を参照します。ECR 自動認証が有効になっている場合は、次のコマンドを実行します。

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--num-executors 2 \
main.py -v
```

ジョブが完了したら、YARN アプリケーション ID を書き留め、次のコマンドを使用して PySpark ジョブの出力を取得します。

```
yarn logs --applicationId application_id | grep -C2 '\[\'
LogLength:55
LogContents:
[[ 0  1  2  3  4]
 [ 5  6  7  8  9]
 [10 11 12 13 14]]
```

## Amazon ECR での SparkR の使用

次の例では、SparkR Dockerfile にタグを付けて ECR にアップロードします。Dockerfile をアップロードすると、SparkR ジョブを実行し、Amazon ECR の Docker イメージを参照できます。

クラスターを起動したら、SSH を使用してコアノードに接続し、次のコマンドを実行して SparkR Dockerfile の例からローカル Docker イメージを構築します。

まず、ディレクトリと Dockerfile を作成します。

```
mkdir sparkr
vi sparkr/Dockerfile
```

SparkR Dockerfile の内容を貼り付け、次のコマンドを実行して Docker イメージを構築します。

```
sudo docker build -t local/sparkr-example sparkr/
```

ローカルに構築したイメージにタグを付けて Amazon ECR にアップロードし、[123456789123.dkr.ecr.us-east-1.amazonaws.com](https://123456789123.dkr.ecr.us-east-1.amazonaws.com) を実際の Amazon ECR エンドポイントに置き換えます。

```
sudo docker tag local/sparkr-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:sparkr-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:sparkr-example
```

SSH を使用してプライマリノードに接続し、sparkR.R というファイル名で R スクリプトを準備します。sparkR.R ファイルに次の内容を貼り付けます。

```
library(SparkR)
sparkR.session(appName = "R with Spark example", sparkConfig =
  list(spark.some.config.option = "some-value"))

sqlContext <- sparkRSQL.init(spark.sparkContext)
library(randomForest)
# check release notes of randomForest
rfNews()

sparkR.session.stop()
```

Amazon EMR 6.0.0 でジョブを送信するには、Docker イメージの名前を参照します。追加の設定パラメータを定義して、ジョブの実行で必ず Docker がランタイムとして使用されるようにします。Amazon ECR を使用する場合は、YARN\_CONTAINER\_RUNTIME\_DOCKER\_CLIENT\_CONFIG は ECR への認証に使用する認証情報が含まれている config.json ファイルを参照する必要があります。

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:sparkr-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
sparkR.R
```



Amazon EMR 6.1.0 以降でジョブを送信するには、Docker イメージの名前を参照します。ECR 自動認証が有効になっている場合は、次のコマンドを実行します。

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
sparkR.R
```

ジョブが完了したら、YARN アプリケーション ID を書き留め、次のコマンドを使用して SparkR ジョブの出力を取得します。次の例では、randomForest ライブラリ、インストールされているバージョン、およびリリースノートが利用可能であることを確認するためのテストを含めています。

```
yarn logs --applicationId application_id | grep -B4 -A10 "Type rfNews"
randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.
Wishlist (formerly TODO):

* Implement the new scheme of handling classwt in classification.

* Use more compact storage of proximity matrix.

* Allow case weights by using the weights in sampling?

=====
Changes in 4.6-14:
```

## AWS Glue データカタログを Spark SQL のメタストアとして使用する

Amazon EMR リリース 5.8.0 以降を使用して、AWS Glue Data Catalog をメタストアとして使用するように Spark SQL を設定できます。永続的なメタストア、またはさまざまなクラスター、サービス、アプリケーション、あるいは AWS アカウントで共有されるメタストアが必要である場合は、この設定をお勧めします。

AWS Glue は、データの分類、クリーニング、強化、さまざまなデータストア間の信頼性の高い移動を簡単かつ費用対効果の高い方法で実現する、フルマネージド型の抽出、変換、ロード (ETL) サービスです。AWS Glue Data Catalog は、Amazon EMR だけでなく、Amazon RDS、Amazon Redshift、Redshift Spectrum、Athena、および Apache Hive メタストアと互換性のあるすべてのアプリケーションと統合されたメタデータリポジトリを提供します。AWS Glue クローラは、Amazon S3 のソースデータからスキーマを自動的に推測し、関連するメタデータをデータカタログに保存できます。データカタログの詳細については、「[AWS Glue デベロッパーガイド](#)」の「[Glue データカタログの入力](#)」を参照してください。AWS

AWS Glue には個別の料金が適用されます。Data Catalog にメタデータを保存してアクセスするための月額料金、AWS Glue ETL ジョブとクローラランタイムに対して 1 分あたりに課金される時間料金、プロビジョニングされた開発エンドポイントごとに 1 分あたりに課金される時間料金があります。Data Catalog では、最大 100 万個までのオブジェクトを無料で保存できます。100 万を超えるオブジェクトを保存した場合は、100,000 オブジェクトごとに 1 USD が課金されます。Data Catalog 内のオブジェクトは、テーブル、パーティション、またはデータベースです。詳細については、「[Glue 料金表](#)」を参照してください。

#### Important

2017 年 8 月 14 日より前に Amazon Athena または Amazon Redshift Spectrum を使用してテーブルを作成した場合、データベースとテーブルは AWS Glue データカタログとは別の Athena 管理カタログに保存されます。Amazon EMR をこれらのテーブルと統合するには、AWS Glue データカタログにアップグレードする必要があります。詳細については、「[Amazon Athena ユーザーガイド](#)」の「[AWS Glue データカタログへのアップグレード](#)」を参照してください。Amazon Athena

## AWS Glue Data Catalog をメタストアとして指定する

、または Amazon EMR API を使用して AWS Management Console AWS CLI、AWS Glue データカタログをメタストアとして指定できます。CLI または API を使用する場合は、Spark の設定分類を使用してデータカタログを指定します。さらに、Amazon EMR 5.16.0 以降では、設定分類を使用して、別の でデータカタログを指定できます AWS アカウント。コンソールを使用する場合は、[Advanced Options] (詳細オプション) または [Quick Options] (クイックオプション) を使用して、Data Catalog を指定できます。

**Note**

Zeppelin は Spark SQL コンポーネントと共にインストールされるため、AWS Glue Data Catalog を使用するオプションは Zeppelin でも使用できます。

**New console**

新しいコンソールで AWS Glue Data Catalog を Spark メタストアとして指定するには

1. にサインインし AWS Management Console、<https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. 左側のナビゲーションペインの [Amazon EMR on EC2] で、[クラスター] を選択し、[クラスターの作成] を選択します
3. [アプリケーションバンドル] で [Spark] または [カスタム] を選択します。クラスターをカスタマイズする場合は、必ず Zeppelin または Spark をアプリケーションの 1 つとして選択します。
4. [AWS Glue Data Catalog 設定] で [Spark テーブルメタデータに使用] チェックボックスをオンにします。
5. クラスターに適用するその他のオプションを選択します。
6. クラスターを起動するには、[クラスターの作成] を選択します。

**Old console**

古いコンソールで AWS Glue Data Catalog を Spark メタストアとして指定するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Release] で、[emr-5.8.0] 以降を選択します。
4. [Release] で、[Spark] または [Zeppelin] を選択します。
5. [AWS Glue Data Catalog 設定] で [Spark テーブルメタデータに使用] を選択します。
6. 必要に応じてクラスターの他のオプションを選択し、[Next (次へ)] 選択してアプリケーションに適切な他のクラスターオプションを設定します。

## AWS CLI

を使用して AWS Glue データカタログを Spark メタストアとして指定するには AWS CLI

AWS CLI と Amazon EMR API を使用して設定分類を指定する方法の詳細については、「」を参照してください [アプリケーションの設定](#)。

- 次の例に示すように、`spark-hive-site` 分類を使用して `hive.metastore.client.factory.class` の値を指定します。

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
      "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

別の AWS アカウントでデータカタログを指定するには、次の例に示すように `hive.metastore.glue.catalogid` プロパティを追加します。`acct-id` は、データカタログの AWS アカウントに置き換えます。

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
      "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

## IAM アクセス許可

クラスターの EC2 インスタンスプロファイルには、AWS Glue アクションの IAM アクセス許可が必要です。さらに、AWS Glue Data Catalog オブジェクトの暗号化を有効にする場合は、暗号化 AWS KMS key に使用される の暗号化、復号、生成もロールに許可する必要があります。

### AWS Glue アクションのアクセス許可

Amazon EMR のデフォルトの EC2 インスタンスプロファイルを使用する場合、アクションは必要ありません。にアタッチされている AmazonElasticMapReduceforEC2Role マネージドポリシーは、必要なすべての AWS Glue アクション EMR\_EC2\_DefaultRole を許可します。ただし、カスタム EC2 インスタンスプロファイルとアクセス許可を指定する場合は、適切な AWS Glue アクションを設定する必要があります。開始点として AmazonElasticMapReduceforEC2Role 管理ポリシーを使用します。詳細については、「Amazon EMR 管理ガイド」の「[クラスター EC2 インスタンスのサービスロール \(EC2 インスタンスプロファイル\)](#)」を参照してください。

### AWS Glue Data Catalog を暗号化および復号するためのアクセス許可

インスタンスプロファイルには、キーを使用してデータを暗号化および復号するためのアクセス許可が必要です。以下のステートメントが両方とも適用される場合、これらのアクセス許可を設定する必要はありません。

- AWS Glue のマネージドキーを使用して、AWS Glue Data Catalog オブジェクトの暗号化を有効にします。
- AWS Glue データカタログ AWS アカウント と同じ にあるクラスターを使用します。

それ以外の場合は、EC2 インスタンスプロファイルにアタッチされたアクセス許可ポリシーに次のステートメントを追加する必要があります。

```
[
  {
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Action": [
          "kms:Decrypt",
          "kms:Encrypt",
          "kms:GenerateDataKey"
        ]
      }
    ]
  }
]
```

```
        "Resource": "arn:aws:kms:region:acct-  
id:key/12345678-1234-1234-1234-123456789012"  
    }  
]  
}  
]
```

AWS Glue データカタログの暗号化の詳細については、「[AWS Glue デベロッパーガイド](#)」の「[データカタログの暗号化](#)」を参照してください。

## リソースベースのアクセス許可

Amazon EMR で AWS Glue を Hive、Spark、または Presto と組み合わせて使用する場合、AWS Glue は Data Catalog リソースへのアクセスを制御するためのリソースベースのポリシーをサポートします。これらのリソースには、データベース、テーブル、接続、ユーザー定義関数が含まれます。詳細については、「[AWS Glue デベロッパーガイド](#)」の「[AWS Glue リソースポリシー](#)」を参照してください。

リソースベースのポリシーを使用して Amazon EMR 内から AWS Glue へのアクセスを制限する場合、アクセス許可ポリシーで指定するプリンシパルは、クラスターの作成時に指定する EC2 インスタンスプロファイルに関連付けられたロール ARN である必要があります。例えば、カタログにアタッチされたリソースベースのポリシーの場合、次の例に示す形式を使用して、クラスター EC2 インスタンスのデフォルトのサービスロールのロール ARN `Principal`、`EMR_EC2_DefaultRole` をとして指定できます。

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

`acct-id` は AWS Glue アカウント ID とは異なる場合があります。これにより、さまざまなアカウントで EMR クラスターからアクセスできます。異なるアカウントから、複数のプリンシパルを指定できます。

## AWS Glue Data Catalog を使用する場合の考慮事項

Spark でメタストアとして AWS Glue Data Catalog を使用する場合は、次の項目を考慮してください。

- 場所の URI が指定されていないデフォルトデータベースを使用すると、テーブルの作成時に障害が発生します。回避策として、LOCATION の使用時に `s3://EXAMPLE-DOC-BUCKET` 句を使用して CREATE TABLE などのバケットの場所を指定します。または、デフォルトデータベース以外のデータベース内にテーブルを作成します。

- AWS Glue 内からのテーブルの名前変更はサポートされていません。
- LOCATION を指定せずに Hive テーブルを作成すると、テーブルデータは、hive.metastore.warehouse.dir プロパティによって指定された場所に保管されます。デフォルトでは、これは HDFS 内の場所です。別のクラスターがテーブルにアクセスする必要がある場合、テーブルを作成したクラスターに対する適切なアクセス許可がない限り、処理に失敗します。さらに、HDFS ストレージは一時的であるため、クラスターが終了すると、テーブルデータは失われ、テーブルを再作成する必要があります。AWS Glue を使用して Hive Amazon S3 LOCATION を指定することをお勧めします。または、hive-site 設定分類を使用して、hive.metastore.warehouse.dir の Amazon S3 で場所を指定できます。これは、すべての Hive テーブルに適用されます。テーブルが HDFS ロケーションで作成され、テーブルを作成したクラスターがまだ実行されている場合は、AWS Glue 内からテーブルロケーションを Amazon S3 に更新できます。詳細については、「[AWS Glue デベロッパーガイド](#)」の「[Glue コンソールでのテーブルの操作](#)」を参照してください。AWS
- 引用符とアポストロフィを含むパーティション値はサポートされていません(例: PARTITION (owner="Doe 's")).).
- [列統計](#)は、emr-5.31.0 以降でサポートされています。
- [Hive 認可](#)の使用はサポートされていません。代替策として、[AWS Glue リソースベースのポリシー](#)を使用することを検討してください。詳細については、[AWS 「Glue データカタログへの Amazon EMR アクセスにリソースベースのポリシーを使用する」](#)を参照してください。

## Spark の設定

設定分類を使用して、[Amazon EMR 上の Spark](#) を設定できます。設定分類の詳細については「[アプリケーションの設定](#)」を参照してください。

Amazon EMR の Spark の設定分類には次が含まれます。

- **spark - maximizeResourceAllocation** プロパティを true または false に設定します。true の場合、Amazon EMR はクラスターハードウェア設定に基づいて、自動的に spark-defaults プロパティを設定します。詳細については、「[maximizeResourceAllocation の使用](#)」を参照してください。
- **spark-defaults** - spark-defaults.conf ファイルに値を設定します。詳細については、Spark ドキュメントの「[Spark Configuration](#)」を参照してください。
- **spark-env** - spark-env.sh ファイルに値を設定します。詳細については、Spark ドキュメントの「[Environment Variables](#)」を参照してください。



- **spark-hive-site** - Spark の `hive-site.xml` に値を設定します。
- **spark-log4j** — (Amazon EMR リリース 6.7.x 以前) `log4j.properties` ファイルに値を設定します。詳細については Github の [log4j.properties.template](#) ファイルをご覧ください。
- **spark-log4j2** — (Amazon EMR リリース 6.8.0 以降) `log4j2.properties` ファイルに値を設定します。詳細については、Github の [log4j2.properties.template](#) ファイルをご覧ください。
- **spark-metrics** - `metrics.properties` ファイルに値を設定します。設定と詳細については Github の [metrics.properties.template](#) ファイルと Spark ドキュメントの [Metrics](#) をご覧ください。

#### Note

Spark ワークロードを別のプラットフォームから Amazon EMR に移行する場合、カスタム設定を追加する前に、[Amazon EMR で設定される Spark のデフォルト](#) でワークロードをテストすることをお勧めします。ほとんどのお客様は、デフォルト設定でパフォーマンスが向上されます。

## トピック

- [Amazon EMR で設定される Spark のデフォルト](#)
- [Amazon EMR 6.1.0 での Spark ガベージコレクションの設定](#)
- [maximizeResourceAllocation の使用](#)
- [ノード停止の動作設定](#)
- [Spark ThriftServer 環境変数](#)
- [Spark のデフォルト設定の変更](#)
- [Apache Log4j 1.x から Log4j 2.x への移行](#)

## Amazon EMR で設定される Spark のデフォルト

次の表はアプリケーションに影響する `spark-defaults` で Amazon EMR がどのようにデフォルト値を設定するか示しています。



## Amazon EMR で設定される Spark のデフォルト

設定	説明	デフォルト値
<code>spark.executor.memory</code>	エグゼキュタープロセスごとに使用するメモリ量。例: 1g、2g。	この設定は、クラスター内のコアインスタンスタイプとタスクインスタンスタイプによって決まります。
<code>spark.executor.cores</code>	各エグゼキュターに使用するコアの数。	この設定は、クラスター内のコアインスタンスタイプとタスクインスタンスタイプによって決まります。
<code>spark.dynamicAllocation.enabled</code>	true の場合、動的なリソース割り当てを使用して、ワークロードによって、アプリケーションで登録するエグゼキュターの数进行调整します。	true (Amazon EMR 4.4.0 以降を使用)
<code>spark.sql.hive.advancedPartitionPredicatePushdown.enabled</code>	true の場合、Hive メタストアへの高度なパーティション述語プッシュダウンが有効になります。	true
<code>spark.sql.hive.stringLikePartitionPredicatePushdown.enabled</code>	<code>startsWith</code> 、 <code>contains</code> 、および <code>endsWith</code> フィルターを Hive メタストアにプッシュダウンします。	true

**Note**  
Spark Shuffle Service は Amazon EMR によって自動的に設定されます。

**Note**  
Glue は `startsWith`、`contains`、ま

設定	説明	デフォルト値
	<p>たは <code>endsWith</code> の述語プッシュダウンをサポートしていません。Glue メタストアを使用していて、これらの関数の述語プッシュダウンが原因でエラーが発生した場合は、この設定を <code>false</code> にしてください。</p>	

## Amazon EMR 6.1.0 での Spark ガベージコレクションの設定

`spark.driver.extraJavaOptions` と `spark.executor.extraJavaOptions` を使用してカスタムガーベージコレクション構成を設定すると、Amazon EMR 6.1.0 とガーベージコレクション設定で競合が生じるため、Amazon EMR 6.1 でのドライバーまたはエグゼキューター起動が失敗します。Amazon EMR 6.1.0 の場合、デフォルトのガーベージコレクション設定は `spark.driver.defaultJavaOptions` および `spark.executor.defaultJavaOptions` を介して設定されます。この設定は、Amazon EMR 6.1.0 にのみ適用されます。ガーベージコレクションに関連しない JVM オプション (ロギングの設定用のオプションなど (`-verbose:class`)) は、`extraJavaOptions` で設定できます。詳細については、「[Spark application properties](#)」を参照してください。

## `maximizeResourceAllocation` の使用

クラスター内の各ノードでリソースを最大限に使用できるようにエグゼキューターを設定するには、`spark` 設定分類で `maximizeResourceAllocation` を `true` に設定します。`maximizeResourceAllocation` は Amazon EMR に固有のもので、`maximizeResourceAllocation` を有効にすると、Amazon EMR はコアインスタンスグループのインスタンスのエグゼキューターで利用可能な最大のコンピューティングとメモリリソースを計算します。次に、計算された最大値に基づいて対応する `spark-defaults` 設定を設定します。

Amazon EMR は、コアインスタンスフリートのインスタンスタイプに基づいて、エグゼキュターで使用できる最大コンピューティングリソースとメモリアリソースを計算します。各インスタンスフリートはフリート内で異なるインスタンスタイプとサイズを持つことができるため、Amazon EMR が使用するエグゼキュター設定はクラスターにとって最適ではない可能性があるため、最大リソース割り当てを使用するときはデフォルト設定を使用することはお勧めしません。インスタンスフリートクラスターのカスタム設定を行います。

#### Note

HBase などの他の分散アプリケーションでは、クラスターで `maximizeResourceAllocation` オプションを使用しないでください。Amazon EMR は分散アプリケーションにカスタム YARN 設定を使用するため、`maximizeResourceAllocation` と競合し、Spark アプリケーションが失敗する可能性があります。

以下に、`maximizeResourceAllocation` が `true` に設定された Spark 設定分類の例を示します。

```
[
  {
    "Classification": "spark",
    "Properties": {
      "maximizeResourceAllocation": "true"
    }
  }
]
```

設定は `spark-defaults` で設定されます (`maximizeResourceAllocation` が有効になっている場合)

設定	説明	値
<code>spark.default.parallelism</code>	結合などの変換によって返される RDDs のデフォルトのパーティション数。ユーザーが設定しない場合 <code>reduceByKey</code> 、並列化されます。	YARN コンテナで使用可能な CPU コアの数 (2 の倍数)。

設定	説明	値
spark.driver.memory	ドライバープロセスに使用するメモリの量。SparkContext は初期化されます (1g、2g など)。	設定は、クラスター内のスレーブインスタンスタイプに基づいて定義されます。ただし、Spark ドライバーアプリケーションは、プライマリインスタンスまたはいずれかのコアインスタンスで (例えば、YARN クライアントモードとクラスターモードのそれぞれで) 実行されるため、この設定は、これらの 2 つのインスタンスグループ間で、小さい方のインスタンスタイプに基づいて定義されます。
spark.executor.memory	エグゼキュターのプロセスごとに使用するメモリの量 (例: 1g、2g)	設定は、クラスター内のコアおよびタスクインスタンスタイプに基づいて行われます。
spark.executor.cores	各エグゼキュターに使用するコアの数。	設定は、クラスター内のコアおよびタスクインスタンスタイプに基づいて行われます。
spark.executor.instances	エグゼキュターの数。	設定は、クラスター内のコアおよびタスクインスタンスタイプに基づいて行われます。同時に spark.dynamicAllocation.enabled が明示的に true に設定されていない限り設定されません。

## ノード停止の動作設定

Amazon EMR リリース 5.9.0 以降では、Amazon EMR の Spark には、手動によるサイズ変更または自動スケーリングポリシーの要求によるノードの終了を Spark が適切に処理できるようにする機能のセットが含まれています。Amazon EMR では Spark に拒否リストメカニズムを実装しています。これは YARN の停止メカニズムに基づいて構築されています。このメカニズムにより、停止中のノードで新しいタスクがスケジュールされないようにし、同時に、既に実行中のタスクを完了するようにします。さらに、ノードの終了時にシャッフルブロックが失われた場合に、Spark のジョブを迅速に回復するための機能があります。再計算プロセスはすぐに実行され、より少ないステージの再試行でより迅速に再計算を行うように最適化されます。シャッフルブロックが見つからないことにより発生するフェッチのエラーによるジョブの失敗を避けることができます。

### Important

Amazon EMR リリース 5.11.0 に `spark.decommissioning.timeout.threshold` 設定が追加され、スポットインスタンスを使用する場合の Spark の耐障害性が向上しました。これまでのリリースでは、ノードがスポットインスタンスを使用していて、インスタンスが入札価格のために終了した場合、Spark は終了を適切に処理できない場合があります。ジョブが失敗する場合があります、シャッフルの再計算に長い時間がかかる場合があります。このため、スポットインスタンスを使用する場合は、リリース 5.11.0 以降を使用することをお勧めします。

### Spark のノード停止の設定

設定	説明	デフォルト値
<code>spark.blacklist.decommissioning.enabled</code>	true に設定すると、Spark は、YARN で <code>decommissioning</code> 状態にあるノードの拒否リストを作成しません。Spark はそのノードで実行されるエグゼキュターに新しいタスクをスケジュールしません。既に実行中のタスクは完了することができます。	true

設定	説明	デフォルト値
spark.blacklist.decommissioning.timeout	<p>decommissioning 状態にあるノードが拒否リストに載せられる時間。デフォルトではこの値は 1 時間に設定されています。これは <code>yarn.resourcemanager.decommissioning.timeout</code> のデフォルトと同じです。ノードが停止期間全体にわたって拒否リストに載せられるようにするには、この値を <code>yarn.resourcemanager.decommissioning.timeout</code> 以上に設定します。停止のタイムアウトの期限が切れた後、ノードは <code>decommissioned</code> 状態に移行し、Amazon EMR はノードの EC2 インスタンスを終了することができます。タイムアウトの期限が切れた後も実行しているタスクがある場合、それらのタスクは失われるか、または強制終了され、他のノードで実行されているエグゼキューターで再スケジュールされます。</p>	1h

設定	説明	デフォルト値
<code>spark.decommissioning.timeout.threshold</code>	<p>Amazon EMR リリース 5.11.0 以降で使用できます。指定は秒単位です。ノードが停止状態に移行する際、ホストがこの値以下の使用停止時間内に停止される場合、Amazon EMR はノードを拒否リストに追加するだけでなく、ノードが停止状態に移行するのを待たずに、ホストの状態を (<code>spark.resourceManager.cleanupExpiredHost</code> の指定に従って) クリーンアップします。これにより、スポットインスタンスが <code>yarn.resourcemanager.decommissioning.timeout</code> の値にかかわらず 20 秒以内にタイムアウトするため、Spark がスポットインスタンスの終了をより良く処理できます。ただし、他のノードがシャッフルファイルを読み取るための時間が足りない場合があります。</p>	20s
<code>spark.resourceManager.cleanupExpiredHost</code>	<p><code>true</code> に設定すると、Spark は、<code>decommissioned</code> 状態にあるノードのエグゼキューターに保存されているすべてのキャッシュデータとシャッフルブロックの登録を解除します。これにより、復旧プロセスが高速化されます。</p>	<code>true</code>

設定	説明	デフォルト値
<code>spark.stage.attempt.ignoreOnDecommissionFetchFailure</code>	true に設定すると、Spark でのステージの失敗と、停止ノードでの多数のフェッチの失敗によるジョブの失敗を回避できます。decommissioned 状態にあるノードのシャッフルブロックのフェッチの失敗は、連続するフェッチのエラーの最大数にカウントされません。	true

## Spark ThriftServer 環境変数

Spark では、Hive Thrift Server Port 環境変数である `HIVE_SERVER2_THRIFT_PORT` が 10001 に設定されます。

## Spark のデフォルト設定の変更

`spark-defaults` 設定分類を使用して、`spark-defaults.conf` 内のデフォルトを変更するか、`spark` 設定分類の `maximizeResourceAllocation` 設定を使用します。

次の手順では、CLI またはコンソールを使用して設定を変更する方法を示します。

CLI を使用して、`spark.executor.memory` が 2g に設定されたクラスターを作成する

- Amazon S3 に格納されている `myConfig.json` ファイルを参照する次のコマンドを使用して、Spark がインストールされ、`spark.executor.memory` が 2g に設定されたクラスターを作成します。

```
aws emr create-cluster --release-label emr-7.1.0 --applications Name=Spark \
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/mybucket/myfolder/myConfig.json
```



**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

myConfig.json:

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.executor.memory": "2G"
    }
  }
]
```

コンソールを使用して、spark.executor.memory が 2g に設定されたクラスターを作成する

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [Spark] を選択します。
4. [Edit software settings] (ソフトウェア設定の編集) で、[Enter configuration] (設定の入力) を選択したままにしておき、次の設定を入力します。

```
classification=spark-defaults,properties=[spark.executor.memory=2G]
```

5. 他のオプションを選択し、 を選択して、[Create cluster] (クラスターの作成) を選択します。

## を設定するには maximizeResourceAllocation

- Spark がインストールされ、Amazon S3 myConfig.jsonに保存されているファイル を参照して AWS CLI、 を使用して true にmaximizeResourceAllocation設定されたクラスターを作成します。

```
aws emr create-cluster --release-label emr-7.1.0 --applications Name=Spark \  
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2 \  
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://  
s3.amazonaws.com/mybucket/myfolder/myConfig.json
```

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

### myConfig.json:

```
[  
  {  
    "Classification": "spark",  
    "Properties": {  
      "maximizeResourceAllocation": "true"  
    }  
  }  
]
```

### Note

Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## Apache Log4j 1.x から Log4j 2.x への移行

[Apache Spark](#) リリース 3.2.x 以前のリリースでは、レガシー Apache Log4j 1.x とその `log4j.properties` ファイルを使用して Spark プロセスで Log4j を設定します。Apache Spark リリース 3.3.0 以降のリリースでは、Apache Log4j 2.x とその `log4j2.properties` ファイルを使用して Spark プロセスで Log4j を設定します。

6.8.0 より前の Amazon EMR リリースを使用して Apache Spark Log4j を設定した場合、Amazon EMR 6.8.0 以降にアップグレードする前に、従来の `spark-log4j` 設定分類を削除し、`spark-log4j2` 設定分類とキー形式に移行する必要があります。Amazon EMR リリース 6.8.0 以降では、従来の `spark-log4j` 分類によりクラスターの作成が `ValidationException` エラーで失敗します。Log4j の非互換性に関連する障害については課金されませんが、続行するには無効になっている `spark-log4j` 設定分類を削除する必要があります。

Apache Log4j 1.x から Log4j 2.x への移行に関する詳細は、Github の「[Apache Log4j Migration Guide](#)」と「[Spark Log4j 2 Template](#)」を参照してください。

### Note

Amazon EMR では、Apache Spark は「[Apache Log4j Migration Guide](#)」で説明されている.xml ファイルではなく `log4j2.properties` ファイルを使用します。また、Log4j 1.x ブリッジメソッドを使用して Log4j 2.x に変換することはお勧めしません。

## Spark パフォーマンスの最適化

Amazon EMR では、Spark 向けに複数のパフォーマンス最適化機能が用意されています。このトピックでは、それぞれの最適化機能について詳しく説明します。

Spark の設定方法に関する詳細については、「[Spark の設定](#)」を参照してください。

### アダプティブクエリ実行

アダプティブクエリ実行は、ランタイム統計に基づいてクエリプランを再最適化するためのフレームワークです。Amazon EMR 5.30.0 以降、Apache Spark 3 からの次のアダプティブクエリ実行最適化が Spark 2 の Apache Amazon EMR ランタイムで利用できます。

- アダプティブ結合変換
- シャッフルパーティションのアダプティブ結合

## アダプティブ結合変換

アダプティブ結合変換は、クエリステージのランタイムサイズに基づいてオペレーションを broadcast-hash-joins オペレーションに変換 sort-merge-join することで、クエリのパフォーマンスを向上させます。B は、結合の一方の側がすべてのエグゼキュターに出力を効率的にブロードキャストできるほど小さい場合、パフォーマンスが向上する broadcast-hash-joins 傾向があるため、交換をシャッフルして結合の両側をソートする必要がなくなります。アダプティブ結合変換は、Spark が自動的に を実行する場合の範囲を広げます broadcast-hash-joins。

この機能は、デフォルトでご利用になれます。これは、`spark.sql.adaptive.enabled` を `false` に設定することで無効にできます。これにより、アダプティブクエリ実行フレームワークも無効になります。Spark は、結合側の 1 つの broadcast-hash-join ランタイムサイズ統計が を超えていない場合に、`sort-merge-join` を に変換することを決定します。デフォルトでは `10spark.sql.autoBroadcastJoinThreshold,485,760` バイト (10 MiB) です。

### シャッフルパーティションのアダプティブ結合

シャッフルパーティションのアダプティブ結合により、小さな連続するシャッフルパーティションが結合され、小さなタスクが多すぎることによるオーバーヘッドを回避することで、クエリのパフォーマンスが向上します。これにより、事前に初期シャッフルパーティションの数をより高く設定することができ、ランタイムでターゲットサイズまで削減されるため、シャッフルパーティションがより均等に分散される可能性が高まります。

この機能は、`spark.sql.shuffle.partitions` が明示的に設定されていない限り、デフォルトで有効になっています。これは、`spark.sql.adaptive.coalescePartitions.enabled` を `true` に設定することで有効にできます。シャッフルパーティションの初期数とターゲットパーティションのサイズは、それぞれ `spark.sql.adaptive.coalescePartitions.minPartitionNum` プロパティと `spark.sql.adaptive.advisoryPartitionSizeInBytes` プロパティを使用して調整できます。この機能に関連する Spark プロパティの詳細については、以下の表を参照してください。

### Spark アダプティブ結合パーティションのプロパティ

プロパティ	デフォルト値	説明
<code>spark.sql.adaptive.coalescePartitions.enabled</code>	<code>spark.sql.shuffle.partitions</code> が明示的に設定されている場合を除き、 <code>true</code>	<code>true</code> であり、かつ、 <code>spark.sql.adaptive.enabled</code> が <code>true</code> の場合、Spark はターゲットサイズ ( <code>spark.sql.adaptive.advisory</code>

プロパティ	デフォルト値	説明
		PartitionSizeInBytes によって指定) に従って、連続するシャッフルパーティションを結合し、小さいタスクが多くなり過ぎないようにします。
spark.sql.adaptive.advisoryPartitionSizeInBytes	64MB	結合時のシャッフルパーティションのアドバイザリサイズ (バイト単位)。この設定は、spark.sql.adaptive.enabled と spark.sql.adaptive.coalescePartitions.enabled が両方とも true の場合にのみ効果があります。
spark.sql.adaptive.coalescePartitions.minPartitionNum	25	結合後のシャッフルパーティションの最小数。この設定は、spark.sql.adaptive.enabled と spark.sql.adaptive.coalescePartitions.enabled が両方とも true の場合にのみ効果があります。
spark.sql.adaptive.coalescePartitions.initialPartitionNum	1,000	結合前のシャッフルパーティションの初期数。この設定は、spark.sql.adaptive.enabled と spark.sql.adaptive.coalescePartitions.enabled が両方とも true の場合にのみ効果があります。

## ダイナミックパーティションプルーニング

ダイナミックパーティションプルーニングでは、特定のパーティションを特定のクエリを読み取って処理するテーブル内でより正確に選択し、ジョブパフォーマンスを改善します。読み取って処理するデータ量を減らすと、ジョブの実行において大幅な時間の節約になります。Amazon EMR 5.26.0 では、この機能はデフォルトで有効になっています。Amazon EMR 5.24.0 および 5.25.0 では、Spark 内から Spark プロパティ `spark.sql.dynamicPartitionPruning.enabled` を設定して有効にできます。また、クラスターの作成時に有効にすることもできます。

### Spark ダイナミックパーティションプルーニングのパーティションのプロパティ

プロパティ	デフォルト値	説明
<code>spark.sql.dynamicPartitionPruning.enabled</code>	true	true の場合、ダイナミックパーティションプルーニングが有効になります。
<code>spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse</code>	true	true の場合、Spark はクエリの実行前に防御チェックを行い、ダイナミックプルーニングフィルターでのブロードキャスト交換の再利用が、ユーザー定義の列指向ルールなどの後続の準備ルールによって破損していないことを確認します。再利用が壊れていて、この config が true の場合、Spark は影響を受けるダイナミックプルーニングフィルターを除去し、パフォーマンスと正確性に関する問題から保護します。ダイナミックプルーニングフィルターのブロードキャスト交換によって、対応する結合オペレーションのブロードキャスト交換とは異なる、整合しない結果が生じると、正確性

プロパティ	デフォルト値	説明
		に関する問題が発生します。この構成ではユーザー定義の列指向ルールによって再利用が破損している場合などのシナリオを回避できるため、これを false に設定する場合は慎重に行ってください。アダプティブクエリ実行が有効になっている場合、ブロードキャストの再利用は常に適用されます。

この最適化は、予定時間に解決される固定述語のプッシュダウンにのみ対応している Spark 2.4.2 の既存の機能を改良するものです。

Spark 2.4.2 での固定述語のプッシュダウンの例を以下に示します。

```
partition_col = 5

partition_col IN (1,3,5)

partition_col between 1 and 3

partition_col = 1 + 3
```

ダイナミックパーティションプルーニングにより、Spark エンジンでランタイムに積極的に情報から推計でき、パーティションが読み取り、安全に削除できます。たとえば、以下のクエリには、すべてのストアの販売合計を含み、リージョンごとに分けられた store\_sales テーブルと、各国のリージョンのマッピングを含む store\_regions テーブルという、2 つのテーブルが含まれます。テーブルには世界中に分散しているストアに関するデータが含まれていますが、北米のデータのみをクエリします。

```
select ss.quarter, ss.region, ss.store, ss.total_sales
from store_sales ss, store_regions sr
where ss.region = sr.region and sr.country = 'North America'
```

ダイナミックパーティションプルーニングを使用しない場合、このクエリはサブクエリ結果に一致する地域のサブセットをフィルタリングする前に、すべての地域のデータを読み取ります。ダイナミックパーティションプルーニングを使用すると、このクエリはサブクエリで返された地域のパーティションのみを読み取り処理します。このため、ストレージのデータの読み取りや履歴の処理が少なくなり、時間とリソースが削減されます。

## スカラサブクエリの平坦化

この最適化では、同じテーブルにスカラサブクエリのあるクエリのパフォーマンスを向上させます。Amazon EMR 5.26.0 では、この機能はデフォルトで有効になっています。Amazon EMR 5.24.0 および 5.25.0 では、Spark プロパティ `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` を Spark 内で設定して有効にできます。また、クラスターの作成時に有効にすることもできます。このプロパティを `true` に設定すると、可能な場合に、クエリオプティマイザが同じ関係式を使用する集約スカラサブクエリを平坦化します。スカラサブクエリは、サブクエリ内に存在する述語を集約関数にプッシュし、リレーションごとにすべての集約関数を使用して1つの集約を実行することによって平坦化されます。

以下は、この最適化を行うメリットがあるクエリの例です。

```
select (select avg(age) from students          /* Subquery 1 */
       where age between 5 and 10) as group1,
       (select avg(age) from students          /* Subquery 2 */
       where age between 10 and 15) as group2,
       (select avg(age) from students          /* Subquery 3 */
       where age between 15 and 20) as group3
```

最適化により、以前のクエリが次のように書き換えられます。

```
select c1 as group1, c2 as group2, c3 as group3
from (select avg (if(age between 5 and 10, age, null)) as c1,
            avg (if(age between 10 and 15, age, null)) as c2,
            avg (if(age between 15 and 20, age, null)) as c3 from students);
```

書き換えられたクエリは一度だけ student テーブルを読み取ります。また、3つのサブクエリの述語が avg 関数にプッシュされます。



## INTERSECT 前に DISTINCT

この最適化では、INTERSECT を使用する場合の結合を最適化します。Amazon EMR 5.26.0 では、この機能はデフォルトで有効になっています。Amazon EMR 5.24.0 および 5.25.0 では、Spark プロパティ `spark.sql.optimizer.distinctBeforeIntersect.enabled` を Spark 内で設定して有効にできます。また、クラスターの作成時に有効にすることもできます。INTERSECT を使用したクエリは、左準結合を使用するように自動変換されます。このプロパティが true に設定されている場合、DISTINCT 演算子が BroadcastHashJoin の代わりに左半結合をにできることを検出すると、クエリオプティマイザは DISTINCT 演算子を INTERSECT の子にプッシュします SortMergeJoin。

以下は、この最適化を行うメリットがあるクエリの例です。

```
(select item.brand brand from store_sales, item
  where store_sales.item_id = item.item_id)
intersect
(select item.brand cs_brand from catalog_sales, item
  where catalog_sales.item_id = item.item_id)
```

このプロパティが有効になっていないと

`spark.sql.optimizer.distinctBeforeIntersect.enabled`、クエリは以下のように書き換えられます。

```
select distinct brand from
  (select item.brand brand from store_sales, item
    where store_sales.item_id = item.item_id)
left semi join
  (select item.brand cs_brand from catalog_sales, item
    where catalog_sales.item_id = item.item_id)
on brand <=> cs_brand
```

このプロパティを有効にすると `spark.sql.optimizer.distinctBeforeIntersect.enabled`、クエリは以下のように書き換えられます。

```
select brand from
  (select distinct item.brand brand from store_sales, item
    where store_sales.item_id = item.item_id)
left semi join
  (select distinct item.brand cs_brand from catalog_sales, item
    where catalog_sales.item_id = item.item_id)
on brand <=> cs_brand
```

## ブルームフィルター結合

この最適化では、結合の一方の側の値から生成された [Bloom フィルター](#) を使用して、結合のもう一方の側を事前にフィルタリングすることで、一部の結合のパフォーマンスを向上させます。Amazon EMR 5.26.0 では、この機能はデフォルトで有効になっています。Amazon EMR 5.25.0 では、Spark 内で Spark プロパティ `spark.sql.bloomFilterJoin.enabled` を `true` に設定してこの機能を有効にできます。また、クラスターを作成するときにもこの機能を有効にできます。

以下は、Bloom フィルターを使用することによってメリットがあるクエリの例です。

```
select count(*)
from sales, item
where sales.item_id = item.id
and item.category in (1, 10, 16)
```

この機能を有効にすると、Bloom フィルターが、クエリの対象となるカテゴリのセットに属するカテゴリを持つすべてのアイテム ID から構築されます。sales テーブルのスキャン中に Bloom フィルターを使用して、Bloom フィルターで定義されたセット内に確実に含まれていないアイテムの売上を洗い出します。こうすることで、特定されたこれらの売上をできるだけ早い段階で除外できます。

## 結合順序の最適化

この最適化では、フィルターを持つテーブルを含む結合の順序を変更することでクエリのパフォーマンスを向上させることができます。Amazon EMR 5.26.0 では、この機能はデフォルトで有効になっています。Amazon EMR 5.25.0 では、Spark 設定パラメータ `spark.sql.optimizer.sizeBasedJoinReorder.enabled` を `true` に設定することで、この機能を有効にすることができます。Spark のデフォルトでは、クエリにリストされている順に、テーブルと左から右に結合するようになっています。この手法では、フィルターを含む小さい結合を最初に実行し、コストの高い結合を後で行うというチャンスを逃してしまいます。

このクエリの例では、ある国のすべての店舗に返品されたすべての商品が報告されます。結合順序を最適化しなかった場合、Spark は、大きなテーブル 2 つ (`store_sales` と `store_returns`) を最初に結合した後、これを `store` と結合し、最後に `item` と結合します。

```
select ss.item_value, sr.return_date, s.name, i.desc,
from store_sales ss, store_returns sr, store s, item i
where ss.id = sr.id and ss.store_id = s.id and ss.item_id = i.id
and s.country = 'USA'
```

結合順序を最適化した場合、Spark は、まず `store_sales` を `store` と結合します。 `store` にはフィルターがあり、 `store_returns` および `broadcastable` よりも小さいためです。次に `store_returns`、最後に `item` と結合します。 `item` にフィルターがあり、ブロードキャストが可能な場合は、結合順序の変更の対象となるため、 `store_sales` が `store` と結合され、続いて `item`、最後に `store_returns` と結合されます。

## Spark 結果フラグメントキャッシュ

Amazon EMR 6.6.0 以降には、結果フラグメントを自動的にキャッシュするオプションの Spark 結果フラグメントキャッシュ機能が含まれています。これらの結果フラグメントは、選択した Amazon S3 バケットに保存されているクエリのサブツリーからの結果の一部です。保存されたクエリの結果フラグメントは、その後のクエリ実行で再利用されるため、クエリが高速になります。

結果フラグメント キャッシュは、Spark SQL クエリを分析し、対象となる結果フラグメントを指定した S3 の場所にキャッシュします。それ以降のクエリ実行では、使用可能なクエリの結果フラグメントが自動的に検出され、S3 から取得されます。結果フラグメントキャッシュは結果セットキャッシュとは異なります。結果セットキャッシュでは、それ以降のクエリがキャッシュから結果を返すには元のクエリと完全に一致する必要があります。結果フラグメントキャッシュは、データの静的なサブセットを繰り返しターゲットとするクエリに使用すると、パフォーマンスが大幅に向上します。

次のクエリを考えてみましょう。このクエリでは、2022 年までの注文がカウントされます。

```
select
  l_returnflag,
  l_linestatus,
  count(*) as count_order
from
  lineitem
where
  l_shipdate <= current_date
  and year(l_shipdate) == '2022'
group by
  l_returnflag,
  l_linestatus
```

時間の経過に従い、このクエリを毎日実行して、その年の総売上をレポートする必要があります。結果フラグメントキャッシュを使用しないと、1 年のすべての日の結果を毎日再計算する必要があります。クエリは時間が経つにつれて遅くなり、365 日分の結果をすべて再計算する必要がある年末に最も遅くなります。

結果フラグメントキャッシュを有効にすると、キャッシュから 1 年のうちのそれ以前のすべての日の結果を使用します。毎日、この機能が再計算する必要があるのは 1 日分の結果のみです。この機能は結果フラグメントを計算した後で、フラグメントをキャッシュします。その結果、キャッシュを有効にしたクエリ時間は短く、それ以降の各クエリで一定に保たれます。

## Spark 結果フラグメントキャッシュを有効にする

Spark 結果フラグメントキャッシュを有効にするには、以下の手順を実行します。

1. Amazon S3 にキャッシュバケットを作成し、EMRFS の読み取り/書き込みアクセスを許可します。詳細については、「[Amazon S3 内の EMRFS データへのアクセスを許可する](#)」を参照してください。
2. Amazon EMR Spark 設定を行い、この機能を有効にします。

```
spark.subResultCache.enabled = true
spark.subResultCache.fs.root.path = s3://DOC-EXAMPLE-BUCKET/cache_dir/
```

3. バケットの S3 ライフサイクル管理を有効にして、キャッシュファイルを自動的にクリーンアップします。
4. オプションで、`reductionRationThreshold` および `maxBufferSize` プロパティを設定して、機能をさらに調整します。

```
spark.sql.subResultCache.reductionRatioThreshold
spark.sql.subResultCache.maxBufferSize
```

## 結果フラグメントキャッシュを使用するときの考慮事項

再計算するのではなく、Amazon S3 に既にキャッシュされている結果を使用する場合のコスト削減は、同じキャッシュ結果を使用できる回数に応じて大きくなります。大きなテーブルスキャンの後にフィルターやハッシュ集計を行うクエリでは、結果のサイズを 8 倍以上減らすことができれば (つまり、入力サイズ:結果の比率が 8:1 以上)、この機能のメリットは最も大きくなります。入力と結果の削減率が大きいほど、コスト面でのメリットも大きくなります。結果を生成するコストが Amazon S3 から結果を取得するコストよりも大きい限り、削減率は小さくても、テーブルスキャンとフィルターまたは集計の間に高価な計算ステップが含まれるクエリにもメリットがあります。デフォルトでは、結果フラグメントキャッシュは、削減率が少なくとも 8:1 になると検出された場合にのみ有効になります。

キャッシュされた結果をクエリが繰り返し再利用する場合、この機能のメリットが最も大きくなります。ローリングウィンドウクエリや増分ウィンドウクエリが良い例です。例えば、30 日間のローリングウィンドウクエリを既に 29 日間実行した場合、元の入カソースからターゲットデータの 30 分の 1 を取得するだけで、過去 29 日間のキャッシュされた結果フラグメントが使用されます。増分ウィンドウクエリは、ウィンドウの開始が固定されているため、さらにメリットがあります。クエリを呼び出すたびに、入カソースからの読み取りが必要になる処理の割合は小さくなります。

結果フラグメントキャッシュを使用する場合のその他の考慮事項は次のとおりです。

- 同じクエリフラグメントで、同じデータを対象としないクエリは、キャッシュヒットレートが低くなるため、この機能のメリットはありません。
- 削減率が低く、コストのかかる計算ステップを含まないクエリでは、結果がキャッシュされ、最初に処理したときとほぼ同じくらいの読み取りコストが発生します。
- 最初のクエリでは、キャッシュへの書き込みコストがかかるため、常に軽微なリグレーションが発生します。
- 結果フラグメントキャッシュ機能は Parquet ファイルでのみ機能します。その他のフェイル形式はサポートされていません。
- 結果フラグメントキャッシュ機能のバッファは、ファイル分割サイズが 128 MB 以上のスキャンのキャッシュのみを試みます。デフォルトの Spark 構成では、スキャンサイズ (スキャン対象のすべてのファイルの合計サイズ) をエグゼキューターコアの数で割った値が 128 MB 未満の場合、結果フラグメントキャッシュは無効になります。以下に示す Spark 構成のいずれかが設定されている場合、ファイル分割サイズは次のようになります。

```
min(maxPartitionBytes, max(openCostInBytes, scan size / minPartitionNum))
```

- `spark.sql.leafNodeDefaultParallelism` (デフォルト値は `spark.default.parallelism`)
- `spark.sql.files.minPartitionNum` (デフォルト値は `spark.sql.leafNodeDefaultParallelism`)
- `spark.sql.files.openCostInBytes`
- `spark.sql.files.maxPartitionBytes`
- 結果フラグメントキャッシュ機能は RDD パーティションの詳細度でキャッシュします。前述の削減率 (デフォルトは 8:1) は、RDD パーティションごとに評価されます。RDD あたりの削減率が 8:1 より大きいワークロードと 8:1 未満のワークロードの両方が含まれる場合、RDD あたりの削減率が常に 8:1 未満のワークロードよりも、パフォーマンス上のメリットが小さくなることがあります。
- 結果フラグメントキャッシュ機能は、キャッシュされる各 RDD パーティションにデフォルトで 16 MB の書き込みバッファを使用します。RDD パーティションごとに 16 MB を超える容量が

キャッシュされる場合、書き込みが不可能であると判断するコストにより、パフォーマンスが低下する可能性があります。

- デフォルトでは、結果フラグメントキャッシュは 8:1 未満の削減率では RDD パーティションの結果をキャッシュしようとせず、書き込みバッファを 16 MB に制限しますが、これらの値は両方とも以下の設定で調整できます。

```
spark.sql.subResultCache.reductionRatioThreshold (default: 8.0)
spark.sql.subResultCache.maxBufferSize (default: 16MB, max: 64MB)
```

- 同じ Amazon EMR リリースを使用する複数のクラスターは、同じキャッシュ場所を共有できません。結果の正確性を確保するため、結果フラグメントキャッシュでは、異なるリリースの Amazon EMR で書き込まれたキャッシュ結果は使用されません。
- 結果フラグメントキャッシュは、Spark Streaming のユースケース、または RecordServer、Apache Ranger、AWS Lake Formation が使用されているときに自動的に無効になります。
- 結果フラグメントキャッシュの読み取り/書き込みでは、EMRFS と Amazon S3 バケットが使用されます。CSE/SSE S3/SSE KMS 暗号化がサポートされています。

## Apache Spark 用の Nvidia RAPIDS アクセラレータの使用

Amazon EMR リリース 6.2.0 以降では、Nvidia による [Apache Spark プラグイン用 RAPIDS](#) アクセラレーターを使用し、EC2 グラフィックス処理ユニット (GPU) インスタンスタイプを使用して Spark を加速できます。RAPIDS Accelerator は、コードを変更せずに Apache Spark 3.0 データサイエンスパイプラインの GPU を加速させ、データ処理とモデルのトレーニングをスピードアップしつつ、インフラストラクチャコストを大幅に低減します。

以下のセクションでは、Spark 用 Spark-RAPIDS プラグインを使用するように EMR クラスターを構成する方法について説明します。

### インスタンスタイプの選択

Spark 用の Nvidia Spark-RAPIDS プラグインを使用するには、コアインスタンスグループとタスクインスタンスグループが、Spark-RAPIDS の [ハードウェア要件](#) を満たす EC2 GPU インスタンスタイプを使用する必要があります。Amazon EMR でサポートされる GPU インスタンスタイプの完全なリストについては、「Amazon EMR 管理ガイド」の「[サポートされるインスタンスタイプ](#)」を参照してください。プライマリインスタンスグループのインスタンスタイプは GPU タイプまたは非 GPU タイプのいずれかになりますが、ARM インスタンスタイプはサポートされていません。

## クラスターのアプリケーション設定のセットアップ

1. Amazon EMR が新しいクラスターにプラグインをインストールできるようにする

プラグインをインストールするには、クラスターの作成時に次の設定を指定します。

```
{
  "Classification": "spark",
  "Properties": {
    "enableSparkRapids": "true"
  }
}
```

2. GPU を使用するように YARN を設定する

YARN での GPU の使用法の詳細については、Apache Hadoop ドキュメントの「[YARN で GPU を使用する](#)」を参照してください。以下の例は、Amazon EMR 6.x および 7.x リリースの YARN 設定のサンプルを示しています。

Amazon EMR 7.x

Amazon EMR 7.x の YARN 設定の例

```
{
  "Classification": "yarn-site",
  "Properties": {
    "yarn.nodemanager.resource-plugins": "yarn.io/gpu",
    "yarn.resource-types": "yarn.io/gpu",
    "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices": "auto",
    "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables": "/usr/
bin",
    "yarn.nodemanager.linux-container-executor.cgroups.mount": "true",
    "yarn.nodemanager.linux-container-executor.cgroups.mount-path": "/spark-
rapids-cgroup",
    "yarn.nodemanager.linux-container-executor.cgroups.hierarchy": "yarn",
    "yarn.nodemanager.container-
executor.class": "org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
  }
}, {
  "Classification": "container-executor",
  "Properties": {
  },
}
```



```
"Configurations":[
  {
    "Classification":"gpu",
    "Properties":{"
      "module.enabled":"true"
    }
  },
  {
    "Classification":"cgroups",
    "Properties":{"
      "root":"/spark-rapids-cgroup",
      "yarn-hierarchy":"yarn"
    }
  }
]
```

## Amazon EMR 6.x

### Amazon EMR 6.x の YARN 設定の例

```
{
  "Classification":"yarn-site",
  "Properties":{"
    "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
    "yarn.resource-types":"yarn.io/gpu",
    "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
    "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
    "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
    "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
    "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
    "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
  }
},{
  "Classification":"container-executor",
  "Properties":{"
  },
  "Configurations":[
    {
      "Classification":"gpu",
```



```

        "Properties":{
            "module.enabled":"true"
        }
    },
    {
        "Classification":"cgroups",
        "Properties":{
            "root":"/sys/fs/cgroup",
            "yarn-hierarchy":"yarn"
        }
    }
]
}

```

### 3. RAPIDS を使用するように Spark を設定する

Spark が RAPIDS プラグインを使用できるようにするために必要な設定は次のとおりです。

```

{
  "Classification":"spark-defaults",
  "Properties":{
    "spark.plugins":"com.nvidia.spark.SQLPlugin",
    "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/gpu/getGpusResources.sh",
    "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native"
  }
}

```

XGBoost ドキュメントの [XGBoost4J-Spark ライブラリ](#) は、クラスターで Spark RAPIDS プラグインが有効になっている場合も使用できます。次の設定を使用して、XGBoost を Spark ジョブと統合できます。

```

{
  "Classification":"spark-defaults",
  "Properties":{
    "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-spark_3.0-1.4.2-0.3.0.jar"
  }
}

```

GPU が加速された EMR クラスターの調整に使用できるその他の Spark 設定については、Nvidia.github.io ドキュメントの「[Rapids Accelerator for Apache Spark tuning guide](#)」を参照してください。

#### 4. YARN キャパシティスケジューラを設定する

DominantResourceCalculator は、GPU のスケジューリングと分離が有効になるように設定する必要があります。詳細については、Apache Hadoop ドキュメントの「[Using GPU on YARN](#)」を参照してください。

```
{
  "Classification": "capacity-scheduler",
  "Properties": {
    "yarn.scheduler.capacity.resource-calculator": "org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
  }
}
```

#### 5. JSON ファイルを作成してすべての設定を含める

Spark クラスターに RAPIDS プラグインを使用するための設定が含まれた JSON ファイルを作成できます。このファイルは、後でクラスターの起動時に指定します。

ファイルはローカルまたは S3 に保存できます。クラスターにアプリケーション設定を指定する方法の詳細については、[アプリケーションの設定](#) を参照してください。

以下のサンプルファイルをテンプレートとして使用して、独自の設定を構築してください。

Amazon EMR 7.x

#### Amazon EMR 7.x 用のサンプル `my-configurations.json` ファイル

```
[
  {
    "Classification": "spark",
    "Properties": {
      "enableSparkRapids": "true"
    }
  },
  {
    "Classification": "yarn-site",
    "Properties": {
      "yarn.nodemanager.resource-plugins": "yarn.io/gpu",

```

```

        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},
{
    "Classification":"container-executor",
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
        },
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/spark-rapids-cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
},
{
    "Classification":"spark-defaults",
    "Properties":{
        "spark.plugins":"com.nvidia.spark.SQLPlugin",
        "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
        "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",

```

```

        "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
        "spark.rapids.sql.concurrentGpuTasks":"1",
        "spark.executor.resource.gpu.amount":"1",
        "spark.executor.cores":"2",
        "spark.task.cpus":"1",
        "spark.task.resource.gpu.amount":"0.5",
        "spark.rapids.memory.pinnedPool.size":"0",
        "spark.executor.memoryOverhead":"2G",
        "spark.locality.wait":"0s",
        "spark.sql.shuffle.partitions":"200",
        "spark.sql.files.maxPartitionBytes":"512m"
    }
},
{
    "Classification":"capacity-scheduler",
    "Properties":{
        "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
    }
}
]

```

## Amazon EMR 6.x

### Amazon EMR 6.x 用のサンプル **my-configurations.json** ファイル

```

[
  {
    "Classification":"spark",
    "Properties":{
      "enableSparkRapids":"true"
    }
  },
  {
    "Classification":"yarn-site",
    "Properties":{
      "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
      "yarn.resource-types":"yarn.io/gpu",
      "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
      "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
      "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",

```

```

        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},
{
    "Classification":"container-executor",
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
        },
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/sys/fs/cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
},
{
    "Classification":"spark-defaults",
    "Properties":{
        "spark.plugins":"com.nvidia.spark.SQLPlugin",
        "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
        "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
        "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
        "spark.rapids.sql.concurrentGpuTasks":"1",
        "spark.executor.resource.gpu.amount":"1",
        "spark.executor.cores":"2",
        "spark.task.cpus":"1",

```

```
        "spark.task.resource.gpu.amount": "0.5",
        "spark.rapids.memory.pinnedPool.size": "0",
        "spark.executor.memoryOverhead": "2G",
        "spark.locality.wait": "0s",
        "spark.sql.shuffle.partitions": "200",
        "spark.sql.files.maxPartitionBytes": "512m"
    }
},
{
    "Classification": "capacity-scheduler",
    "Properties": {
        "yarn.scheduler.capacity.resource-calculator": "org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
    }
}
]
```

## クラスターのブートストラップアクションを追加する

クラスターの作成時にブートストラップアクションスクリプトを指定する方法の詳細については、「Amazon EMR 管理ガイド」の「[ブートストラップアクションの基本](#)」を参照してください。

以下のサンプルスクリプトは、Amazon EMR 6.x と 7.x のブートストラップアクションファイルを作成する方法を示しています。

### Amazon EMR 7.x

#### Amazon EMR 7.x 用のサンプル `my-bootstrap-action.sh` ファイル

YARN を使用して Amazon EMR 7.x リリースの GPU リソースを管理するには、クラスターに手動で CGroup v1 をマウントする必要があります。この例のように、ブートストラップアクションスクリプトとしてこれを実行できます。

```
#!/bin/bash
set -ex

sudo mkdir -p /spark-rapids-cgroup/devices
sudo mount -t cgroup -o devices cgroupv1-devices /spark-rapids-cgroup/devices
sudo chmod a+rwx -R /spark-rapids-cgroup
```

## Amazon EMR 6.x

### Amazon EMR 6.x 用のサンプル `my-bootstrap-action.sh` ファイル

Amazon EMR 6.x リリースでは、クラスターで YARN CGroup へのアクセス権限を開く必要があります。これは、この例のようにブートストラップアクションスクリプトを使用して実行できます。

```
#!/bin/bash
set -ex

sudo chmod a+rwx -R /sys/fs/cgroup/cpu,cpuacct
sudo chmod a+rwx -R /sys/fs/cgroup/devices
```

## クラスターを起動する


最後のステップは、上記のクラスター設定を使用してクラスターを起動することです。以下は、Amazon EMR CLI からクラスターを起動するコマンドの例です。

```
aws emr create-cluster \
--release-label emr-7.1.0 \
--applications Name=Hadoop Name=Spark \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=my-key-pair,InstanceProfile=EMR_EC2_DefaultRole \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m4.4xlarge \
                    InstanceGroupType=CORE,InstanceCount=1,InstanceType=g4dn.2xlarge \
                    InstanceGroupType=TASK,InstanceCount=1,InstanceType=g4dn.2xlarge \
--configurations file:///my-configurations.json \
--bootstrap-actions Name='My Spark Rapids Bootstrap action',Path=s3://my-bucket/my-
bootstrap-action.sh
```

## Spark シェルにアクセスする

Spark シェルは Scala REPL (Read-Eval-Print-Loop) をベースにしています。このシェルを使用すると、Spark プログラムをインタラクティブに作成し、作業をフレームワークに送信できます。Spark シェルには、SSH を使用してプライマリノードに接続し、`spark-shell` を呼び出すことでアクセスできます。プライマリノードへの接続の詳細については、「Amazon EMR 管理ガイド」の

「[Connect to the primary node using SSH](#)」を参照してください。次の例では、Amazon S3 に格納された Apache HTTP Server アクセスログを使用します。

 Note

これらの例にあるバケットは、米国東部 (バージニア北部) にアクセスできるクライアントが使用できます。

デフォルトでは、Spark シェルは という独自の [SparkContext](#) オブジェクトを作成します `sc`。このコンテキストは、REPL 内で必要な場合に使用できます。 `sqlContext` はシェルでも使用でき、です [HiveContext](#)。

Example Spark を使用して Amazon S3 に格納されたファイルにおける文字列の出現回数をカウントする

この例では、Amazon S3 に保存されているテキスト ファイルを読み取るために `sc` を使用します。

```
scala> sc
res0: org.apache.spark.SparkContext = org.apache.spark.SparkContext@404721db

scala> val textFile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/
impressions/dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

Spark は、`textFile` と関連付けられた [データ構造](#) を作成します。次に、この例ではログファイル内の文字列 "cartoonnetwork.com" を含む行の数をカウントします。

```
scala> val linesWithCartoonNetwork = textFile.filter(line =>
  line.contains("cartoonnetwork.com")).count()
linesWithCartoonNetwork: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[2] at
  filter at <console>:23
<snip>
<Spark program runs>
scala> linesWithCartoonNetwork
res2: Long = 9
```



Example Python ベースの Spark シェルを使用して Amazon S3 に格納されたファイルにおける文字列の出現回数をカウントする

Spark には、Python ベースのシェルである pyspark も用意されており、Python で記述された Spark プログラムのプロトタイプを作成するために使用できます。と同様に spark-shell、プライマリノードで pyspark を呼び出します。同じ [SparkContext](#) オブジェクトもあります。

```
>>> sc
<pyspark.context.SparkContext object at 0x7fe7e659fa50>
>>> textfile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/impressions/
dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

Spark は、textFile と関連付けられた [データ構造](#) を作成します。次に、この例ではログファイル内の文字列 "cartoonnetwork.com" を含む行の数をカウントします。

```
>>> linesWithCartoonNetwork = textfile.filter(lambda line: "cartoonnetwork.com" in
line).count()
15/06/04 17:12:22 INFO lzo.GPLNativeCodeLoader: Loaded native gpl library from the
embedded binaries
15/06/04 17:12:22 INFO lzo.LzoCodec: Successfully loaded & initialized native-lzo
library [hadoop-lzo rev EXAMPLE]
15/06/04 17:12:23 INFO fs.EmrFileSystem: Consistency disabled, using
com.amazon.ws.emr.hadoop.fs.s3n.S3NativeFileSystem as filesystem implementation
<snip>
<Spark program continues>
>>> linesWithCartoonNetwork
9
```

## 機械学習に Amazon SageMaker Spark を使用する

Amazon EMR リリース 5.11.0 以降を使用する場合、aws-sagemaker-spark-sdk コンポーネントが Spark と共にインストールされます。このコンポーネントは、Amazon SageMaker Spark および Spark と [Amazon SageMaker](#) との統合に関連する依存関係をインストールします。Amazon SageMaker Spark を使用して、Amazon SageMaker ステージを使用して Spark 機械学習 (ML) パイプラインを構築できます。詳細については、「[Amazon SageMaker デベロッパーガイド](#)」の「[での Amazon Spark README GitHub](#)」および「[Amazon での Apache Spark の使用 SageMaker](#)」を参照してください。 SageMaker

# Spark アプリケーションを作成する

[Spark](#) アプリケーションは、Scala、Java、または Python で記述できます。Apache Spark ドキュメントの「[Spark の例](#)」には、Spark アプリケーションの例がいくつかあります。ネイティブにサポートされる以下の 3 つのアプリケーションには、Estimating Pi の例が示されています。`$SPARK_HOME/examples` および `examples` で完全な例を表示することもできます。[GitHub](#)。Spark 用の JAR を構築する方法の詳細については、Apache Spark ドキュメントの「[クイック スタート](#)」を参照してください。

## Scala

Scala の互換性の問題を回避するために、Amazon EMR クラスター用の Spark アプリケーションをコンパイルするときに、正しい Scala バージョンの Spark の依存関係を使用することをお勧めします。使用すべき Scala バージョンは、クラスターにインストールされている Spark のバージョンによって異なります。例えば、Amazon EMR リリース 5.30.1 は Scala 2.11 でビルドされた Spark 2.4.5 を使用します。クラスターが Amazon EMR リリース 5.30.1 を使用している場合は、Scala 2.11 の Spark の依存関係を使用します。Spark で使用されている Scala バージョンの詳細については、「[Apache Spark ドキュメント](#)」を参照してください。

```
package org.apache.spark.examples
import scala.math.random
import org.apache.spark._

/** Computes an approximation to pi */
object SparkPi {
  def main(args: Array[String]) {
    val conf = new SparkConf().setAppName("Spark Pi")
    val spark = new SparkContext(conf)
    val slices = if (args.length > 0) args(0).toInt else 2
    val n = math.min(100000L * slices, Int.MaxValue).toInt // avoid overflow
    val count = spark.parallelize(1 until n, slices).map { i =>
      val x = random * 2 - 1
      val y = random * 2 - 1
      if (x*x + y*y < 1) 1 else 0
    }.reduce(_ + _)
    println("Pi is roughly " + 4.0 * count / n)
    spark.stop()
  }
}
```

# Java

```
package org.apache.spark.examples;

import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaRDD;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.api.java.function.Function;
import org.apache.spark.api.java.function.Function2;

import java.util.ArrayList;
import java.util.List;

/**
 * Computes an approximation to pi
 * Usage: JavaSparkPi [slices]
 */
public final class JavaSparkPi {

    public static void main(String[] args) throws Exception {
        SparkConf sparkConf = new SparkConf().setAppName("JavaSparkPi");
        JavaSparkContext jsc = new JavaSparkContext(sparkConf);

        int slices = (args.length == 1) ? Integer.parseInt(args[0]) : 2;
        int n = 100000 * slices;
        List<Integer> l = new ArrayList<Integer>(n);
        for (int i = 0; i < n; i++) {
            l.add(i);
        }

        JavaRDD<Integer> dataSet = jsc.parallelize(l, slices);

        int count = dataSet.map(new Function<Integer, Integer>() {
            @Override
            public Integer call(Integer integer) {
                double x = Math.random() * 2 - 1;
                double y = Math.random() * 2 - 1;
                return (x * x + y * y < 1) ? 1 : 0;
            }
        }).reduce(new Function2<Integer, Integer, Integer>() {
            @Override
            public Integer call(Integer integer, Integer integer2) {
                return integer + integer2;
            }
        });
    }
}
```

```
    }
  });

  System.out.println("Pi is roughly " + 4.0 * count / n);

  jsc.stop();
}
}
```

## Python

```
import argparse
import logging
from operator import add
from random import random

from pyspark.sql import SparkSession

logger = logging.getLogger(__name__)
logging.basicConfig(level=logging.INFO, format="%(levelname)s: %(message)s")

def calculate_pi(partitions, output_uri):
    """
    Calculates pi by testing a large number of random numbers against a unit circle
    inscribed inside a square. The trials are partitioned so they can be run in
    parallel on cluster instances.

    :param partitions: The number of partitions to use for the calculation.
    :param output_uri: The URI where the output is written, typically an Amazon S3
                       bucket, such as 's3://example-bucket/pi-calc'.
    """

    def calculate_hit(_):
        x = random() * 2 - 1
        y = random() * 2 - 1
        return 1 if x**2 + y**2 < 1 else 0

    tries = 100000 * partitions
    logger.info(
        "Calculating pi with a total of %s tries in %s partitions.", tries, partitions
    )
```

```
with SparkSession.builder.appName("My PyPi").getOrCreate() as spark:
    hits = (
        spark.sparkContext.parallelize(range(tries), partitions)
        .map(calculate_hit)
        .reduce(add)
    )
    pi = 4.0 * hits / tries
    logger.info("%s tries and %s hits gives pi estimate of %s.", tries, hits, pi)
    if output_uri is not None:
        df = spark.createDataFrame([(tries, hits, pi)], ["tries", "hits", "pi"])
        df.write.mode("overwrite").json(output_uri)

if __name__ == "__main__":
    parser = argparse.ArgumentParser()
    parser.add_argument(
        "--partitions",
        default=2,
        type=int,
        help="The number of parallel partitions to use when calculating pi.",
    )
    parser.add_argument(
        "--output_uri", help="The URI where output is saved, typically an S3 bucket."
    )
    args = parser.parse_args()

    calculate_pi(args.partitions, args.output_uri)
```

## Amazon S3 で Spark のパフォーマンスを向上させる

Amazon EMR には、Spark を使用して Amazon S3 に保存されたデータのクエリ、読み取り、および書き込みを行うときのパフォーマンスを最適化するのに役立つ機能が用意されています。

[S3 Select](#) では、Amazon S3 に処理を「プッシュダウン」することで一部のアプリケーションの CSV および JSON ファイルのクエリパフォーマンスを向上させることができます。

EMRFS S3-optimizedコミッターは、[OutputCommitter](#) クラスの代替手段です。このコミッターは EMRFS のマルチパートアップロード機能を使用して DataFrames、Spark SQL、およびデータセットを使用して Amazon S3 に Parquet ファイルを書き込む際のパフォーマンスを向上させます。

## トピック

- [S3 Select と Spark を使用してクエリパフォーマンスを向上させる](#)
- [EMRFS S3 向けに最適化されたコミッターを使用する](#)
- [EMRFS S3 向けに最適化されたコミットプロトコルを使用する](#)
- [EMRFS で Amazon S3 リクエストを再試行する](#)

## S3 Select と Spark を使用してクエリパフォーマンスを向上させる

Amazon EMR リリース 5.17.0 以降では、Amazon EMR での Spark を使用した [S3 Select](#) を使用できます。S3 Select では、アプリケーションはオブジェクトに含まれるデータのサブセットのみを取得できます。Amazon EMR では、大量のデータセットをフィルタリングして処理する計算作業をクラスターから Amazon S3 に「プッシュダウン」することにより、一部のアプリケーションのパフォーマンスを高めることができます。また、Amazon EMR と Amazon S3 の間で転送されるデータの量も削減されます。

S3 Select は、`s3selectCSV` および `s3selectJSON` 値を使用してデータ形式を指定する CSV ファイルと JSON ファイルでサポートされます。詳細な説明と例については、「[コードで S3 Select を指定する](#)」を参照してください。

### S3 Select が使用するアプリケーションに適しているかどうかを確認する

S3 Select が使用するアプリケーションに適しているかどうかを確認するために、S3 Select を使用した場合と使用しない場合のアプリケーションのベンチマークを行うことをお勧めします。

アプリケーションが S3 Select を使用する候補となるかどうかを判断するには、次のガイドラインを使用します。

- クエリは元のデータセットの半分以上を除外する。
- Amazon S3 と Amazon EMR クラスター間のネットワーク接続は、転送速度と使用可能な帯域幅が良好です。Amazon S3 は、HTTP 応答を圧縮しないため、応答サイズは圧縮された入力ファイルと比較して増加する可能性があります。

### 考慮事項と制約事項

- お客様が用意した暗号化キーを使用した Amazon S3 サーバー側の暗号化 (SSE-C) とクライアント側の暗号化はサポートされていません。

- AllowQuotedRecordDelimiters プロパティはサポートされていません。このプロパティを指定した場合、クエリは失敗します。
- UTF-8 形式の CSV ファイルと JSON ファイルのみがサポートされています。複数行の CSV はサポートされません。
- 圧縮されていないファイルか gzip ファイルのみがサポートされます。
- nanValue、positiveInf、negativeInf などの Spark CSV と JSON のオプションや破損した記録に関連するオプション (failfast および dropmalformed モードなど) はサポートされません。
- 小数でのカンマ (,) の使用はサポートされません。たとえば、10,000 はサポートされませんが、10000 はサポートされます。
- 最後の行のコメント文字はサポートされていません。
- ファイルの末尾にある空の行は処理されません。
- 次のフィルターは Amazon S3 にプッシュダウンされません。
  - COUNT() や SUM() などの集計関数。
  - 属性を CAST() するフィルター。例えば CAST(stringColumn as INT) = 1 です。
  - オブジェクトの属性、または複雑な属性を持つフィルター。例えば intArray[1] = 1, objectColumn.objectNumber = 1 です。
  - 値がリテラル値ではないフィルター。例えば、intColumn1 = intColumn2。
  - 確認された制限のある [S3 Select がサポートするデータ型](#)のみがサポートされます。

## コードで S3 Select を指定する

次の例は、Scala、SQL、R、および `spark` を使用して CSV に S3 Select を指定する方法を示しています。PySpark。JSON 用の S3 Select も同じ方法で使用できます。オプション、デフォルト値、および制限の一覧については、「[オプション](#)」を参照してください。

### PySpark

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional
  .load("s3://path/to/my/datafiles")
```

## R

```
read.df("s3://path/to/my/datafiles", "s3selectCSV", schema, header = "true",
        delimiter = "\t")
```

## Scala

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional. Examples:
  // .options(Map("quote" -> "\"", "header" -> "true")) or
  // .option("quote", "\"").option("header", "true")
  .load("s3://path/to/my/datafiles")
```

## SQL

```
CREATE TEMPORARY VIEW MyView (number INT, name STRING) USING s3selectCSV OPTIONS
  (path "s3://path/to/my/datafiles", header "true", delimiter "\t")
```

## オプション

s3selectCSV と s3selectJSON を使用するときは、以下のオプションを利用できます。指定しない場合はデフォルト値が使用されます。

## S3selectCSV のオプション

オプション	デフォルト値	使用方法
compression	"none"	圧縮が使用されているかどうかを示します。"none" 以外にサポートされる設定は "gzip" のみです。
delimiter	","	フィールドの区切り記号を指定します。
quote	"\""	引用文字を指定します。空の文字列は指定できず、文字列



オプション	デフォルト値	使用方法
		を空にすると XML が不正というエラーが表示されます。
escape	'\\'	エスケープ文字を指定します。
header	"false"	"false" はヘッダーがないことを指定します。"true" はヘッダーが最初の行にあることを指定します。最初の行のヘッダーのみがサポートされ、ヘッダーの前の空の行はサポートされません。
コメント	"#"	コメント文字を指定します。コメントインジケータを無効にすることはできません。つまり、\u0000 という値はサポートされません。
nullValue	""	

### S3selectJSON のオプション

オプション	デフォルト値	使用方法
compression	"none"	圧縮が使用されているかどうかを示します。"none" 以外にサポートされる設定は "gzip" のみです。
multiline	"false"	"false" は JSON が S3 Select LINES 形式であることを指定し、入力データの各行に 1 つの JSON オブジェクトが含まれていることを意味し

オプション	デフォルト値	使用方法
		<p>ます。"true" は JSON が S3 Select DOCUMENT 形式であることを指定し、入力データの複数の行に JSON オブジェクトをまたがらせられることを意味します。</p>

## EMRFS S3 向けに最適化されたコミッターを使用する

EMRFS S3-optimizedコミッターは、EMRFS の使用時に Amazon S3 にファイルを書き込むために最適化された代替 [OutputCommitter](#) 実装です。EMRFS S3 向けに最適化されたコミッターは、ジョブおよびタスクコミットフェーズ中に Amazon S3 で行われるリストオペレーションと名前変更オペレーションを回避することにより、アプリケーションのパフォーマンスを向上させることができます。コミッターは Amazon EMR リリース 5.19.0 以降で使用でき、Amazon EMR 5.20.0 以降ではデフォルトで有効になっています。コミッターは、Spark SQL、DataFrames、またはデータセットを使用する Spark ジョブに使用されます。Amazon EMR 6.4.0 以降では、Parquet、ORC、テキストベースの形式 (CSV と JSON を含む) など、一般的なあらゆる形式にこのコミッターを使用できます。Amazon EMR 6.4.0 より前のリリースでは、Parquet 形式のみがサポートされています。コミッターが使用されない状況があります。詳細については、「[EMRFS S3 向けに最適化されたコミッターの要件](#)」を参照してください。

### トピック

- [EMRFS S3 向けに最適化されたコミッターの要件](#)
- [EMRFS S3 向けに最適化されたコミッターとマルチパートアップロード](#)
- [ジョブの調整に関する考慮事項](#)
- [Amazon EMR 5.19.0 の EMRFS S3 向けに最適化されたコミッターを有効にする](#)

## EMRFS S3 向けに最適化されたコミッターの要件

以下の条件が満たされる場合に、EMRFS S3 向けに最適化されたコミッターが使用されます。

- Spark SQL DataFrames またはデータセットを使用して Amazon S3 にファイルを書き込む Spark ジョブを実行します。Amazon EMR 6.4.0 以降では、Parquet、ORC、テキストベースの形式

(CSV と JSON を含む) など、一般的なあらゆる形式にこのコミッターを使用できます。Amazon EMR 6.4.0 より前のリリースでは、Parquet 形式のみがサポートされています。

- Amazon EMR でマルチパートアップロードが有効になっている。これがデフォルトです。詳細については、「[EMRFS S3 向けに最適化されたコミッターとマルチパートアップロード](#)」を参照してください。
- Spark の組み込みファイル形式のサポートが使用されます。組み込みファイル形式のサポートは以下の状況で使用されます。
  - Hive メタストアテーブルの場合、Parquet テーブルに対して `spark.sql.hive.convertMetastoreParquet` が true に設定される場合、または、Amazon EMR 6.4.0 以降の Orc テーブルに対して `spark.sql.hive.convertMetastoreOrc` が true に設定される場合。これらはデフォルトの設定です。
  - ジョブによってファイル形式のデータソースまたはテーブルに書き込まれる場合。例えば、ターゲットテーブルが USING parquet 句で作成される場合などです。
  - ジョブでパーティション分割されていない Hive メタストア Parquet テーブルに書き込む場合。Spark の組み込み Parquet サポートはパーティション分割された Hive テーブルをサポートしていません。これは既知の制限です。詳細については、「[Apache Spark SQL, DataFrames and Datasets Guide](#)」の「[Hive メタストアの Parquet テーブル変換](#)」を参照してください。
- デフォルトのパーティションの場所 (`${table_location}/k1=v1/k2=v2/` など) に書き込む Spark ジョブオペレーションでコミッターが使用される。ジョブオペレーションによってカスタムのパーティション場所書き込まれる場合、例えば、カスタムのパーティション場所が ALTER TABLE SQL コマンドを使用して設定されている場合、コミッターは使用されません。
- Spark で以下の値を使用する。
  - `spark.sql.parquet.fs.optimized.committer.optimization-enabled` プロパティを true に設定する必要があります。これは、Amazon EMR 5.20.0 以降でのデフォルト設定です。Amazon EMR 5.19.0 では、デフォルト値は false です。この値の設定については、「[Amazon EMR 5.19.0 の EMRFS S3 向けに最適化されたコミッターを有効にする](#)」を参照してください。
  - パーティション分割されていない Hive メタストアテーブルに書き込む場合、Parquet と Orc ファイル形式のみがサポートされます。パーティション分割されていない Parquet Hive メタストアテーブルに書き込む true 場合は、`spark.sql.hive.convertMetastoreParquet` を true に設定する必要があります。パーティション分割されていない Orc Hive メタストアテーブルに書き込む true 場合は、`spark.sql.hive.convertMetastoreOrc` を true に設定する必要があります。これらはデフォルトの設定です。

- `spark.sql.parquet.output.committer.class` を `com.amazon.emr.committer.EmrOptimizedSparkSqlParquetOutputCommitter` に設定する必要があります。これはデフォルトの設定です。
- `spark.sql.sources.commitProtocolClass` は `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol` または `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol` に設定する必要があります。Amazon EMR 5.x シリーズバージョン 5.30.0 以降、および Amazon EMR 6.x シリーズバージョン 6.2.0 以降の場合、`org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol` がデフォルト設定です。それよりも前の Amazon EMR バージョンの場合、`org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol` がデフォルト設定です。
- Spark ジョブでパーティション分割された Parquet データセットを動的パーティション列で上書きする場合は、`partitionOverwriteMode` 書き込みオプションと `spark.sql.sources.partitionOverwriteMode` を `static` に設定する必要があります。これはデフォルトの設定です。

#### Note

`partitionOverwriteMode` 書き込みオプションは Spark 2.4.0 で導入されました。Amazon EMR リリース 5.19.0 に含まれている Spark バージョン 2.3.2 では、`spark.sql.sources.partitionOverwriteMode` プロパティを設定します。

## EMRFS S3 用に最適化されたコミッターが使用されていない場合

一般的に、EMRFS S3 用に最適化されたコミッターは次の状況では、使用されません。

状況	コミッターが使われない理由
HDFS に書き込む場合	コミッターは、EMRFS を使用した Amazon S3 への書き込みのみをサポートします。
S3A ファイルシステムを使用する場合	コミッターは EMRFS のみをサポートします。

状況	コミッターが使われない理由
MapReduce または Spark の RDD API を使用する場合	コミッターは、SparkSQL、DataFrame、またはデータセット APIs の使用のみをサポートします。

以下の Scala の例では、いくつかの追加の状況を示しています。EMRFS S3 向けに最適化されたコミッターを全体に使用しないもの (最初の例) と、部分的に使用しないもの (2 番目の例) です。

#### Example - 動的パーティション上書きモード

以下の Scala の例では、別のコミットアルゴリズムを使用するように Spark に指示しています。これでは、EMRFS S3 向けに最適化されたコミッターはまったく使用されません。このコードは、データを書き込むパーティションのみを上書きするように `partitionOverwriteMode` プロパティを `dynamic` に設定します。次に、動的パーティション列が `partitionBy` で指定され、書き込みモードが `overwrite` に設定されます。

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")
  .option("partitionOverwriteMode", "dynamic")
  .partitionBy("dt")
  .parquet("s3://EXAMPLE-DOC-BUCKET/output")
```

EMRFS S3 向けに最適化されたコミッターが使用されないようにするには、3 つの設定をすべて構成する必要があります。これを行うと、Spark は Spark のコミットプロトコルで指定されている別のコミットアルゴリズムを実行します。5.30.0 より前の Amazon EMR 5.x リリースと 6.2.0 より前の Amazon EMR 6.x リリースの場合、コミットプロトコルは Spark のステージングディレクトリを使用します。これは、`.spark-staging` で始まる出力場所に作成された一時ディレクトリです。このアルゴリズムではパーティションディレクトリの名前が順番に変更されるため、パフォーマンスが低下する可能性があります。Amazon EMR リリース 5.30.0 以降および 6.2.0 以降の詳細については、「[EMRFS S3 向けに最適化されたコミットプロトコルを使用する](#)」を参照してください。

Spark 2.4.0 のアルゴリズムは以下の手順に従います。

1. タスク試行により、Spark のステージングディレクトリ (`${outputLocation}/spark-staging-${jobID}/k1=v1/k2=v2/` など) の下のパーティションディレクトリに出力が書き込まれます。

- 書き込まれたパーティションごとに、タスク試行は相対パーティションパス ( $k1=v1/k2=v2$  など) を管理します。
- タスクが正常に完了すると、追跡されたすべての相対パーティションパスがドライバーに渡されます。
- すべてのタスクが完了した後、ジョブのコミットフェーズでは、成功したタスクの試行によって Spark のステージングディレクトリに書き込まれたすべてのパーティションディレクトリが収集されます。ディレクトリツリーの名前変更オペレーションを使用して、これらの各ディレクトリの名前が最終的な出力場所に順番に変更されます。
- ステージングディレクトリがジョブのコミットフェーズの完了前に削除されます。

### Example - カスタムのパーティション場所

この例では、Scala コードは 2 つのパーティションを挿入します。1 つのパーティションはカスタムのパーティション場所を使用します。もう 1 つのパーティションはデフォルトのパーティション場所を使用します。EMRFS S3 向けに最適化されたコミッターは、デフォルトのパーティション場所を使用するパーティションへのタスク出力の書き込みにのみ使用されます。

```
val table = "dataset"
val location = "s3://bucket/table"

spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
  """)

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
  """)

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
```

```
.withColumn("dt",  
  when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))  
.write.insertInto(table)
```

Scala コードは以下の Amazon S3 オブジェクトを作成します。

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet  
custom_${folder}$  
table/_SUCCESS  
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet  
table/dt=2019-01-29_${folder}$  
table_${folder}$
```

カスタムのパーティション場所に書き込むとき、先ほどの例と同様のコミットアルゴリズムが使用されます。先ほどの例と同様、このアルゴリズムでは名前が順番に変更されるため、パフォーマンスが低下する可能性があります。

1. カスタムのパーティション場所に出力を書き込むとき、タスクでは最終的な出力場所に作成される Spark のステージングディレクトリにファイルを書き込みます。ファイルの名前には、ファイルの競合から保護するためのランダムな UUID が含まれます。タスクの試行によって各ファイルが最終的な出力パスと共に追跡されます。
2. タスクが正常に完了すると、ドライバーにファイルとそれらの最終的な出力パスが渡されます。
3. すべてのタスクが完了した後、ジョブのコミットフェーズでは、カスタムのパーティション場所に書き込まれたすべてのファイルの名前が、最終的な出力パスに順番に変更されます。
4. ステージングディレクトリがジョブのコミットフェーズの完了前に削除されます。

## EMRFS S3 向けに最適化されたコミッターとマルチパートアップロード

EMRFS S3 に最適化されたコミッターを使用するには、Amazon EMR のマルチパートアップロードを有効にする必要があります。マルチパートアップロードは、デフォルトで有効になっています。また、必要に応じて再度有効にすることができます。詳細については、「[Amazon EMR 管理ガイド](#)」の「Amazon S3 用のマルチパートアップロードを設定する」を参照してください。

EMRFS S3 向けに最適化されたコミッターは、マルチパートアップロードのトランザクションのような特性を利用して、タスクのコミット時にタスクの試行によって書き込まれたファイルのみがジョブの出力場所に表示されるようにします。この方法でマルチパートアップロードを使用することで、コミッターはデフォルトの FileOutputCommitter アルゴリズムバージョン 2 よりもタスクコミットの



パフォーマンスを向上させます。EMRFS S3 向けに最適化されたコミッターを使用するにあたっては、従来のマルチパートアップロードとのいくつかの重要な違いを考慮する必要があります。

- マルチパートアップロードは常にファイルサイズに関係なく実行されます。これは EMRFS のデフォルトの動作とは異なり、マルチパートアップロードがトリガーされるファイルサイズは `fs.s3n.multipart.uploads.split.size` プロパティで制御されます。
- タスクがコミットされるか中止されるまで、マルチパートアップロードは長期間不完全な状態になります。これは EMRFS のデフォルトの動作とは異なり、タスクで指定されたファイルの書き込みが終了するとマルチパートアップロードが完了します。

このような違いにより、タスクが実行されて Amazon S3 へのデータの書き込みが行われているときに Spark Executor JVM がクラッシュしたり強制終了されたりすると、不完全なマルチパートアップロードが残される可能性が高くなります。そのため、EMRFS S3 に最適化されたコミッターを使用するときは、失敗したマルチパートアップロードの管理に関するベストプラクティスに従ってください。詳細については、「Amazon EMR 管理ガイド」の Amazon S3 バケットの処理に関する [ベストプラクティス](#) を参照してください。

## ジョブの調整に関する考慮事項

EMRFS S3 向けに最適化されたコミッターは、タスクがコミットされるか中止されるまで、タスクの試行によって書き込まれた各ファイルのために少量のメモリを消費します。ほとんどのジョブで消費されるメモリの量は無視できる程度です。多数のファイルを書き込む長時間のタスクを含むジョブの場合、コミッターが消費するメモリが多くなり、Spark エグゼキュターに割り当てられたメモリの調整が必要になることがあります。エグゼキュターのメモリは、`spark.executor.memory` プロパティを使用して調整できます。ガイドラインとして、100,000 個のファイルを書き込む 1 つのタスクでは、一般的に 100 MB のメモリを追加する必要があります。詳細については、Apache Spark Configuration ドキュメントの「[Application Properties](#)」を参照してください。

## Amazon EMR 5.19.0 の EMRFS S3 向けに最適化されたコミッターを有効にする

Amazon EMR 5.19.0 を使用している場合は、クラスターの作成時、または Spark 内から (Amazon EMR を使用しているとき)、`spark.sql.parquet.fs.optimized.committer.optimization-enabled` プロパティを `true` に手動で設定できます。



## EMRFS S3 向けに最適化されたコミッターをクラスターの作成時に有効にする

`spark.sql.parquet.fs.optimized.committer.optimization-enabled` プロパティを `true` に設定するには、`spark-defaults` 設定分類を使用します。詳細については、「[アプリケーションの設定](#)」を参照してください。

## EMRFS S3 向けに最適化されたコミッターを Spark から有効にする

`spark.sql.parquet.fs.optimized.committer.optimization-enabled` は、SparkConf でハードコーディングするか、Spark シェルまたは `spark-submit` および `spark-sql` ツールで `--conf` パラメータとして渡すか、`conf/spark-defaults.conf` によって `true` に設定できます。詳細については、Apache Spark ドキュメントの「[Spark Configuration](#)」を参照してください。

次の例は、`spark-sql` コマンドの実行中にコミッターを有効にする方法を示しています。

```
spark-sql \  
  --conf spark.sql.parquet.fs.optimized.committer.optimization-enabled=true \  
  -e "INSERT OVERWRITE TABLE target_table SELECT * FROM source_table;"
```

## EMRFS S3 向けに最適化されたコミットプロトコルを使用する

EMRFS S3-optimizedコミットプロトコルは、EMRFS の使用時に Spark 動的パーティション上書きを使用して Amazon S3 にファイルを書き込むために最適化された代替[FileCommitProtocol](#)実装です。このプロトコルは、Spark の動的パーティション上書きジョブのコミットフェーズ中の Amazon S3 での名前変更操作を回避することで、アプリケーションのパフォーマンスを向上させます。

[EMRFS S3 向けに最適化されたコミッターを使用する](#) はまた、名前変更操作を回避することでパフォーマンスを向上させることにも注意してください。ただし、動的パーティション上書きの場合は機能しませんが、コミットプロトコルの改善は動的パーティション上書きの場合のみを対象としています。

コミットプロトコルは Amazon EMR リリース 5.30.0 以降で使用でき、6.2.0 以降ではデフォルトで有効になっています。リリース 5.31.0 以降、Amazon EMR では並列処理の改善が追加されました。プロトコルは、Spark SQL、DataFramesまたはデータセットを使用する Spark ジョブに使用されません。コミットプロトコルが使用されない状況があります。詳細については、「[EMRFS S3 向けに最適化されたコミットプロトコルの要件](#)」を参照してください。

### トピック

- [EMRFS S3 向けに最適化されたコミットプロトコルの要件](#)
- [EMRFS S3 向けに最適化されたコミットプロトコルとマルチパートアップロード](#)


- [ジョブの調整に関する考慮事項](#)

## EMRFS S3 向けに最適化されたコミットプロトコルの要件

以下の条件が満たされる場合に、EMRFS S3 向けに最適化されたコミットプロトコルが使用されません。

- Spark SQL を使用する Spark ジョブ、またはデータセットを実行して DataFrames、パーティションテーブルを上書きします。
- パーティション上書きモードが dynamic である Spark ジョブを実行する。
- Amazon EMR でマルチパートアップロードが有効になっている。これがデフォルトです。詳細については、「[EMRFS S3 向けに最適化されたコミットプロトコルとマルチパートアップロード](#)」を参照してください。
- EMRFS のファイルシステムキャッシュは有効になっています。これがデフォルトです。fs.s3.impl.disable.cache 設定が false に設定されていることを確認します。
- Spark のビルトインデータソースサポートが使用されています。組み込みデータソースサポートは以下の状況で使用されます。
  - ジョブでビルトインのデータソースまたはテーブルに書き込む場合。
  - ジョブで Hive メタストア Parquet テーブルに書き込む場合。これは spark.sql.hive.convertInsertingPartitionedTable と spark.sql.hive.convertMetastoreParquet が両方とも true に設定されている場合に発生します。これらはデフォルトの設定です。
  - ジョブで Hive メタストア ORC テーブルに書き込む場合。これは、spark.sql.hive.convertInsertingPartitionedTable と spark.sql.hive.convertMetastoreOrc が両方 true に設定されている場合に発生します。これらはデフォルトの設定です。
- デフォルトのパーティションの場所 (`${table_location}/k1=v1/k2=v2/` など) に書き込む Spark ジョブオペレーションでコミットプロトコルが使用される。ジョブオペレーションによってカスタムのパーティション場所に書き込まれる場合、例えば、カスタムのパーティション場所が ALTER TABLE SQL コマンドを使用して設定されている場合、プロトコルは使用されません。
- Spark で以下の値を使用する。
  - spark.sql.sources.commitProtocolClass を org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol に設定する必要があります。これは、Amazon EMR リリース 5.30.0 以降および 6.2.0 以降のデフォルト設定です。

- `partitionOverwriteMode` 書き込みオプションまたは `spark.sql.sources.partitionOverwriteMode` は `dynamic` に設定する必要があります。デフォルトの設定は、`static` です。

 Note

`partitionOverwriteMode` 書き込みオプションは Spark 2.4.0 で導入されました。Amazon EMR リリース 5.19.0 に含まれている Spark バージョン 2.3.2 では、`spark.sql.sources.partitionOverwriteMode` プロパティを設定します。

- Spark ジョブが Hive メタストアの Parquet テーブルに上書きされる場合は、`spark.sql.hive.convertMetastoreParquet`、`spark.sql.hive.convertInsertingParquet` および `spark.sql.hive.convertMetastore.partitionOverwriteMode` を `true` に設定する必要があります。これらはデフォルトの設定です。
- Spark ジョブが Hive メタストアの ORC テーブルに上書きされる場合は、`spark.sql.hive.convertMetastoreOrc`、`spark.sql.hive.convertInsertingPartition` および `spark.sql.hive.convertMetastore.partitionOverwriteMode` を `true` に設定する必要があります。これらはデフォルトの設定です。

### Example - 動的パーティション上書きモード

この Scala の例では、最適化がトリガーされます。まず、`partitionOverwriteMode` プロパティを `dynamic` に設定します。これにより、データを書き込んでいるパーティションのみが上書きされます。次に、動的パーティション列が `partitionBy` で指定され、書き込みモードが `overwrite` に設定されます。

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")           // "overwrite" instead of "insert"
  .option("partitionOverwriteMode", "dynamic") // "dynamic" instead of "static"
  .partitionBy("dt")                       // partitioned data instead of
unpartitioned data
  .parquet("s3://EXAMPLE-DOC-BUCKET/output") // "s3://" to use Amazon EMR file
system, instead of "s3a://" or "hdfs://"
```

## EMRFS S3 向けに最適化されたコミットプロトコルが使用されない場合

一般的に、EMRFS S3 向けに最適化されたコミットプロトコルは、オープンソースのデフォルトの Spark SQL コミットプロトコル

`org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol` と同じように機能します。次の状況では、最適化は行われません。

状況	コミットプロトコルが使用されない理由
HDFS に書き込む場合	コミットプロトコルは、EMRFS を使用した Amazon S3 への書き込みのみをサポートします。
S3A ファイルシステムを使用する場合	コミットプロトコルは EMRFS のみをサポートします。
MapReduce または Spark の RDD API を使用する場合	コミットプロトコルは、SparkSQL、DataFrame、またはデータセット APIs の使用のみをサポートします。
動的パーティションの上書きがトリガーされない場合	コミットプロトコルは、動的なパーティション上書きの場合のみを最適化します。その他の場合については、「 <a href="#">EMRFS S3 向けに最適化されたコミッターを使用する</a> 」を参照してください。

次の Scala の例は、EMRFS S3 向けに最適化されたコミットプロトコルが `SQLHadoopMapReduceCommitProtocol` に委任するその他の状況を示しています。

## Example - カスタムパーティションの場所を使用する動的パーティション上書きモード

この例では、Scala プログラムは動的パーティション上書きモードで 2 つのパーティションを上書きします。1 つのパーティションはカスタムのパーティション場所を使用します。もう 1 つのパーティションはデフォルトのパーティション場所を使用します。EMRFS S3 向けに最適化されたコミットプロトコルは、デフォルトのパーティション場所を使用するパーティションのみを改善します。

```
val table = "dataset"
val inputView = "tempView"
val location = "s3://bucket/table"
```

```
spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
  """)

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
  """)

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .createTempView(inputView)

// Set partition overwrite mode to 'dynamic'
spark.sql(s"SET spark.sql.sources.partitionOverwriteMode=dynamic")

spark.sql(s"INSERT OVERWRITE TABLE $table SELECT * FROM $inputView")
```

Scala コードは以下の Amazon S3 オブジェクトを作成します。

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
custom_${folder$}
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_${folder$}
table_${folder$}
```

#### Note

以前の Spark バージョンでは、カスタムパーティションの場所に書き込むと、データが失われる可能性があります。この例では、パーティション dt='2019-01-28' が失われます。

詳細については「[SPARK-35106](#)」を参照してください。Amazon EMR リリース 5.33.0 以降 (6.0.x および 6.1.x は除く) では、この問題は修正されています。

カスタムのパーティション場所書き込むとき、先ほどの例と同様のコミットアルゴリズムが使用されます。先ほどの例と同様、このアルゴリズムでは名前が順番に変更されるため、パフォーマンスが低下する可能性があります。

Spark 2.4.0 のアルゴリズムは以下の手順に従います。

1. カスタムのパーティション場所に出力を書き込むとき、タスクでは最終的な出力場所に作成される Spark のステージングディレクトリにファイルを書き込みます。ファイルの名前には、ファイルの競合から保護するためのランダムな UUID が含まれます。タスクの試行によって各ファイルが最終的な出力パスと共に追跡されます。
2. タスクが正常に完了すると、ドライバーにファイルとそれらの最終的な出力パスが渡されます。
3. すべてのタスクが完了した後、ジョブのコミットフェーズでは、カスタムのパーティション場所書き込まれたすべてのファイルの名前が、最終的な出力パスに順番に変更されます。
4. ステージングディレクトリがジョブのコミットフェーズの完了前に削除されます。

## EMRFS S3 向けに最適化されたコミットプロトコルとマルチパートアップロード

EMRFS S3 向けに最適化されたコミットプロトコルで動的パーティション上書きの最適化を使用するには、Amazon EMR でマルチパートアップロードを有効にする必要があります。マルチパートアップロードは、デフォルトで有効になっています。また、必要に応じて再度有効にすることができます。詳細については、「[Amazon EMR 管理ガイド](#)」の「Amazon S3 用のマルチパートアップロードを設定する」を参照してください。

動的パーティション上書きの間、EMRFS S3 向けに最適化されたコミットプロトコルは、マルチパートアップロードのトランザクションのような特性を利用して、ジョブのコミット時にタスクの試行によって書き込まれたファイルのみがジョブの出力場所に表示されるようにします。このようにマルチパートアップロードを使用することにより、コミットプロトコルはジョブコミットのパフォーマンスをデフォルトの SQLHadoopMapReduceCommitProtocol よりも向上させます。EMRFS S3 向けに最適化されたコミットプロトコルを使用するにあたっては、従来のマルチパートアップロードとのいくつかの重要な違いを考慮する必要があります。

- マルチパートアップロードは常にファイルサイズに関係なく実行されます。これは EMRFS のデフォルトの動作とは異なり、マルチパートアップロードがトリガーされるファイルサイズは `fs.s3n.multipart.uploads.split.size` プロパティで制御されます。
- タスクがコミットされるか中止されるまで、マルチパートアップロードは長期間不完全な状態になります。これは EMRFS のデフォルトの動作とは異なり、タスクで指定されたファイルの書き込みが終了するとマルチパートアップロードが完了します。

このような違いにより、タスクが実行されて Amazon S3 へのデータの書き込みが行われているときに Spark Executor JVM がクラッシュしたり強制終了されたりする、またはジョブが実行しているときに Spark Driver JVM がクラッシュしたり強制終了されたりすると、不完全なマルチパートアップロードが残される可能性が高くなります。そのため、EMRFS S3 に最適化されたコミットプロトコルを使用するときは、失敗したマルチパートアップロードの管理に関するベストプラクティスに従ってください。詳細については、「Amazon EMR 管理ガイド」の Amazon S3 バケットの処理に関する [ベストプラクティス](#) を参照してください。

## ジョブの調整に関する考慮事項

Spark エグゼキュターで、EMRFS S3 向けに最適化されたコミットプロトコルは、タスクがコミットされるか中止されるまで、タスクの試行によって書き込まれた各ファイルのために少量のメモリを消費します。ほとんどのジョブで消費されるメモリの量は無視できる程度です。

Spark ドライバーで、EMRFS S3 向けに最適化されたコミットプロトコルには、ジョブがコミットまたは中止されるまで、コミットされた各ファイルのメタデータ情報を保存するためのメモリが必要です。ほとんどのジョブでは、Spark ドライバーのデフォルトのメモリ設定はごくわずかです。

多数のファイルを書き込む長時間のタスクを含むジョブの場合、コミットプロトコルが消費するメモリが多くなり、Spark、特に Spark エグゼキュターに割り当てられたメモリの調整が必要になることがあります。Spark ドライバーの `spark.driver.memory` プロパティと Spark エグゼキュターの `spark.executor.memory` プロパティを使用してメモリを調整できます。ガイドラインとして、100,000 個のファイルを書き込む 1 つのタスクでは、一般的に 100 MB のメモリを追加する必要があります。詳細については、Apache Spark Configuration ドキュメントの「[Application Properties](#)」を参照してください。

## EMRFS で Amazon S3 リクエストを再試行する

このトピックでは、EMRFS を使用して Amazon S3 にリクエストを行うときに使用できる再試行戦略について説明します。リクエストレートが上昇すると、S3 は新しいレートをサポートするためにスケーリングを試みます。このプロセス中、S3 はリクエストをスロットルし、503 Slow Down エ



ラーを返す場合があります。S3 リクエストの成功率を高めるために、`emrfs-site` 設定でプロパティを設定し、再試行戦略を調整できます。

次の方法で再試行戦略を調整できます。

- デフォルトのエクスポネンシャルバックオフ再試行戦略の再試行上限を増やします。
- `additive-increase/multiplicative-decrease (AIMD)` 再試行戦略を有効にして設定します。AIMD は Amazon EMR リリース 6.4.0 以降でサポートされています。

## デフォルトのエクスポネンシャルバックオフ戦略を使用する

デフォルトでは、EMRFS はエクスポネンシャルバックオフ戦略を使用して Amazon S3 リクエストを再試行します。EMRFS のデフォルトの再試行上限は 15 です。S3 503 Slow Down エラーを回避するために、新しいクラスターを作成するときに、実行中のクラスターで、またはアプリケーションのランタイムで、再試行上限を増やすことができます。

再試行上限を増やすには、`emrfs-site` で `fs.s3.maxRetries` の値を変更する必要があります。以下の設定例では、`fs.s3.maxRetries` をカスタム値の 30 に設定します。

```
[
  {
    "Classification": "emrfs-site",
    "Properties": {
      "fs.s3.maxRetries": "30"
    }
  }
]
```

設定オブジェクトの処理の詳細については、「[アプリケーションの設定](#)」を参照してください。

## AIMD 再試行戦略を使用する

Amazon EMR リリース 6.4.0 以降では、EMRFS は、`additive-increase/multiplicative-decrease (AIMD)` モデルに基づく代替再試行戦略をサポートします。AIMD 再試行戦略は、大規模な Amazon EMR クラスターで作業する場合に特に便利です。

AIMD は、最近の成功したリクエストに関するデータを使用してカスタムリクエストレートを計算します。この戦略では、スロットリングされるリクエストの数と、リクエストごとに必要な合計試行数が減少します。



AIMD 再試行戦略を有効にするには、以下の例のように、`emrfs-site` 設定で `fs.s3.aimd.enabled` プロパティを `true` に設定する必要があります。

```
[
  {
    "Classification": "emrfs-site",
    "Properties": {
      "fs.s3.aimd.enabled": "true"
    }
  }
]
```

設定オブジェクトの処理の詳細については、「[アプリケーションの設定](#)」を参照してください。

## 詳細な AIMD 再試行設定

次の表にリストされているプロパティを設定して、AIMD 再試行戦略を使用するときの再試行の動作を詳細に指定できます。大部分のユースケースで、デフォルト値を使用することをお勧めします。

### 高度な AIMD 再試行戦略のプロパティ

プロパティ	デフォルト値	説明
<code>fs.s3.aimd.increaseIncrement</code>	0.1	連続するリクエストが成功したときにリクエストレートが上昇する速さを制御します。
<code>fs.s3.aimd.reductionFactor</code>	2	Amazon S3 が 503 レスポンスを返すときにリクエストレートが減少する速さを制御します。デフォルト係数の 2 は、リクエストレートを半分に削減します。
<code>fs.s3.aimd.minRate</code>	0.1	リクエストで S3 による持続的なスロットリングが生じる場合のリクエストレートの下限を設定します。
<code>fs.s3.aimd.initialRate</code>	5500	初期リクエストレートを設定します。その後、

プロパティ	デフォルト値	説明
		これは、 <code>fs.s3.aimd.increaseIncrement</code> と <code>fs.s3.aimd.reductionFactor</code> に指定する値に従って変わります。  初期レートは GET リクエストにも使用され、PUT リクエストに比例してスケーリングされます (3500/5500)。
<code>fs.s3.aimd.adjustWindow</code>	2	リクエストレートを調整する頻度を制御し、応答数で測定します。
<code>fs.s3.aimd.maxAttempts</code>	100	リクエストを試行する最大試行回数を設定します。

## Spark ステップを追加する

Amazon EMR ステップを使用すると、EMR クラスターにインストールされた Spark フレームワークに作業を送信できます。詳細については、「Amazon EMR 管理ガイド」の「[ステップ](#)」を参照してください。コンソールと CLI では、`spark-submit` スクリプトをステップとして自動的に実行する Spark アプリケーションステップを使用してこれを行います。API では、ステップを使用し、`spark-submit` を使用して `command-runner.jar` を呼び出します。

Spark へのアプリケーションの送信の詳細については、Apache Spark ドキュメントの「[アプリケーションの送信](#)」を参照してください。

コンソールを使用して Spark ステップを送信する

1. <https://console.aws.amazon.com/emr> で Amazon EMR コンソールを開きます。
2. [Cluster List] で、クラスターの名前を選択します。
3. [Steps (ステップ)] セクションまでスクロールして展開し、[Add step (ステップの追加)] を選択します。

4. [Add Step] ダイアログボックスで、次のようにします。
  - [Step type] では、[Spark application] を選択します。
  - [Name] では、デフォルト名 ( Spark アプリケーション ) を使用するか、または新しい名前を入力します。
  - [Deploy mode] で、[Client] または [Cluster] を選択します。クライアントモードでは、クラスターのプライマリインスタンスでドライバプログラムが起動されますが、クラスターモードではクラスター上でドライバプログラムが起動されます。クライアントモードの場合、ドライバのログ出力はステップログに表示され、クラスターモードの場合、ドライバのログ出力は最初の YARN コンテナのログに表示されます。詳細については、Apache Spark ドキュメントで「[クラスターモードの概要](#)」を参照してください。
  - 目的の Spark-submit オプション を指定します。spark-submit オプションの詳細については、「[spark-submit を使用したアプリケーションの起動](#)」を参照してください。
  - [Application location] で、アプリケーションのローカルまたは S3 URI パスを指定します。
  - [Arguments (引数)] は、フィールドを空のままにします。
  - [Action on failure (失敗時の操作)] では、デフォルトのオプション ([Continue (続行)]) を使用します。
5. 追加を選択します。ステップは、[Pending] というステータスでコンソールに表示されます。
6. ステップが実行されると、ステータスは [Pending (保留中)] から [Running (実行中)]、[Running (実行中)] から [完了済み] に変更されます。ステータスを更新するには、[Actions] (アクション) 列の上にある [Refresh] (更新) アイコンを選択します。
7. ステップの結果は、ロギングを設定している場合、Amazon EMR コンソールの [Cluster Details] (クラスターの詳細) ページの [Log Files] (ログファイル) にあるステップの横に表示されます。オプションで、クラスターを起動したときに設定したログバケットにステップの情報が記載されています。

を使用して Spark に作業を送信するには AWS CLI

クラスターを作成するか、既存のクラスターで `aws emr add-steps` サブコマンドを使用するとき、ステップを送信します。

1. 次の例に示すように、`create-cluster` を使用します。

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.1.0 --
applications Name=Spark \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10] --use-default-roles
```

代わりに、次の例のように `command-runner.jar` を使用できます。

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.1.0 \
--applications Name=Spark --ec2-attributes KeyName=myKey --instance-type m5.xlarge
--instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Spark Program",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[spark-example,SparkPi,10] --use-default-
roles
```

**Note**

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

2. または、既に実行中のクラスターにステップを追加します。 `add-steps` を使用します。

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps
Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10]
```

代わりに、次の例のように `command-runner.jar` を使用できます。

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Spark Program",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[spark-example,SparkPi,10]
```

SDK for Java を使用して作業を Spark に送信するには

1. 次の例は、Java を使用した Spark のあるクラスターにステップを追加する方法を示しています。

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);

StepFactory stepFactory = new StepFactory();
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig sparkStepConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("spark-submit", "--executor-memory", "1g", "--class", "org.apache.spark.examples.SparkPi", "/usr/lib/spark/examples/jars/spark-examples.jar", "10");

StepConfig sparkStep = new StepConfig()
    .withName("Spark Step")
    .withActionOnFailure("CONTINUE")
    .withHadoopJarStep(sparkStepConf);

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

2. ステップの結果は、ステップのログを調べることで確認します。これは、ステップを選択し、ステップを選択し、ログファイルでまたはを選択して、で AWS Management Console 実行できます stderr。 stdout 利用可能なログを表示するには、[View Logs] を選択します。

## Spark のデフォルト設定を上書きする

Spark のデフォルト設定値をアプリケーションごとに上書きすることができます。これは、ステップを使用してアプリケーションを送信するとき (原則としてオプションが `spark-submit` に渡されます) に行うことができます。たとえば、`spark.executor.memory` を変更することで、実行プログラムのプロセスに割り当てられたメモリを変更することができます。次のような引数を含む `--executor-memory` スイッチを指定します。

```
spark-submit --executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/spark/examples/jars/spark-examples.jar 10
```

同様に、`--executor-cores` と `--driver-memory` を調整できます。ステップでは、次の引数をステップに指定できます。

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/spark/examples/jars/spark-examples.jar 10
```

さらに、`--conf` オプションを使用して、組み込みスイッチがない設定を調整することもできます。調整可能な他の設定の詳細については、Apache Spark ドキュメントの「[Spark プロパティの動的なロード](#)」を参照してください。

## Spark アプリケーション履歴を表示する

コンソールでクラスターの詳細ページの [Application user interfaces] (アプリケーションユーザーインターフェイス) タブを使用して、Spark、YARN アプリケーション、および Tez UI の詳細を表示できます。Amazon EMR のアプリケーションユーザーインターフェイス (UI) を使用すると、アクティブなジョブとジョブ履歴のトラブルシューティングや分析を簡単に行うことができます。

詳細については、「Amazon EMR 管理ガイド」の「[アプリケーションの履歴を表示する](#)」を参照してください。

## Spark ウェブ UI にアクセスする

Spark ウェブ UIs 「[クラスターへの接続](#)」セクションで SSH トンネルを作成するか、プロキシを作成します。次に、クラスター ResourceManager の YARN に移動します。アプリケーションの UI 追跡の下のリンクを選択します。アプリケーションが実行されている場合は、が表示されます ApplicationMaster。ドライバーがどこにあるかにかかわらず、ここからアプリケーションマスターのウェブ UI のポート 20888 に移動できます。YARN クライアントモードで実行する場

合、ドライバーはクラスターのプライマリノードにあるかもしれませんが。アプリケーションを YARN クラスターモードで実行している場合、ドライバーはクラスター上のアプリケーションの ApplicationMaster にあります。アプリケーションが終了すると、履歴が表示され、EMR クラスターのプライマリノードの 18080 にある Spark HistoryServer UI ポート番号が表示されます。これは、すでに完了したアプリケーション用です。http://*master-public-dns-name*:18080/ で Spark HistoryServer UI に直接移動することもできます。

Amazon EMR リリース 5.25.0 以降では、SSH 接続を介してウェブプロキシを設定しなくても、コンソールから Spark 履歴サーバー UI にアクセスできます。詳細については、「[永続アプリケーションユーザーインターフェイスの表示](#)」を参照してください。

## Spark 構造化ストリーミング Amazon Kinesis Data Streams コネクタの使用

Amazon EMR リリース 7.1.0 以降では、Spark 構造化ストリーミング Amazon Kinesis Data Streams コネクタがリリースイメージに含まれています。このコネクタを使用すると、Amazon EMR で Spark を使用して、Amazon Kinesis Data Streams に保存されているデータを処理できます。コネクタは、コンシューマータイプ GetRecords (共有スループット) と SubscribeToShard (拡張ファンアウト) の両方をサポートします。この統合はに基づいています [spark-sql-kinesis-connector](#)。コネクタの使用を開始する方法の詳細については、[README](#) を参照してください。

次の例は、コネクタを使用して Amazon EMR で Spark アプリケーションを起動する方法を示しています。

```
spark-submit my_kinesis_streaming_script.py
```

## Amazon EMR での Apache Spark 用の Amazon Redshift インテグレーションの使用

Amazon EMR リリース 6.4.0 以降では、すべてのリリースイメージに、[Apache Spark](#) と Amazon Redshift の間のコネクタが含まれています。このコネクタを使用すると、Amazon EMR で Spark を使用して Amazon Redshift に保存されているデータを処理できます。Amazon EMR リリース 6.4.0 から 6.8.0 では、インテグレーションは [spark-redshift オープンソースコネクタ](#) に基づいています。Amazon EMR リリース 6.9.0 以降では、[Apache Spark の Amazon Redshift インテグレーション](#) がコミュニティバージョンからネイティブインテグレーションに移行されました。

### トピック

- [Amazon Redshift integration for Apache Spark を使用した Spark アプリケーションの起動](#)
- [Apache Spark 用の Amazon Redshift インテグレーションを使用した認証](#)
- [Amazon Redshift に対する読み書き](#)
- [Spark コネクタを使用する際の考慮事項と制限事項](#)

## Amazon Redshift integration for Apache Spark を使用した Spark アプリケーションの起動

Amazon EMR リリース 6.4 から 6.9 では、`--jars` または `--packages` オプションを使用して、次の JAR ファイルのうちどのファイルを使用するかを指定する必要があります。`--jars` オプションは、ローカル、HDFS、または HTTP/S を使用して保存される依存関係を指定します。`--jars` オプションでサポートされている他のファイルの場所については、Spark ドキュメントの「[Advanced Dependency Management](#)」を参照してください。`--packages` オプションでは、パブリック Maven リポジトリに保存されている依存関係を指定します。

- `spark-redshift.jar`
- `spark-avro.jar`
- `RedshiftJDBC.jar`
- `minimal-json.jar`

Amazon EMR リリース 6.10.0 以降は `minimal-json.jar` 依存関係を必要とせず、デフォルトで他の依存関係を各クラスターに自動的にインストールします。以下の例は、Apache Spark 用の Amazon Redshift インテグレーションを使用して Spark アプリケーションを起動する方法を示しています。

### Amazon EMR 6.10.0 +

次の例は、Amazon EMR リリース 6.10 以降で `spark-redshift` コネクタを使用して Spark アプリケーションを起動する方法を示しています。

```
spark-submit my_script.py
```

### Amazon EMR 6.4.0 - 6.9.x

Amazon EMR リリース 6.4 から 6.9 で `spark-redshift` コネクタを使用して Spark アプリケーションを起動するには、次の例に示すように、`--jars` または `--packages` オプションを使用す



する必要があります。なお、`--jars` オプションに指定されているパスは JAR ファイルのデフォルトのパスであることに注意してください。

```
spark-submit \  
  --jars /usr/share/aws/redshift/jdbc/RedshiftJDBC.jar,/usr/share/aws/redshift/  
spark-redshift/lib/spark-redshift.jar,/usr/share/aws/redshift/spark-redshift/lib/  
spark-avro.jar,/usr/share/aws/redshift/spark-redshift/lib/minimal-json.jar \  
  my_script.py
```

## Apache Spark 用の Amazon Redshift インテグレーションを使用した認証

AWS Secrets Manager を使用して認証情報を取得し、Amazon Redshift に接続する

次のコードサンプルは、AWS Secrets Manager を使用して、Python の Apache Spark 用の PySpark インターフェイスを使用して Amazon Redshift クラスターに接続するための認証情報を取得する方法を示しています。

```
from pyspark.sql import SQLContext  
import boto3  
  
sc = # existing SparkContext  
sql_context = SQLContext(sc)  
  
secretsmanager_client = boto3.client('secretsmanager')  
secret_manager_response = secretsmanager_client.get_secret_value(  
    SecretId='string',  
    VersionId='string',  
    VersionStage='string'  
)  
username = # get username from secret_manager_response  
password = # get password from secret_manager_response  
url = "jdbc:redshift://redshifthost:5439/database?user=" + username + "&password=" +  
password  
  
# Read data from a table  
df = sql_context.read \  
    .format("io.github.spark_redshift_community.spark.redshift") \  
    .option("url", url) \  
    .option("dbtable", "my_table") \  
    .option("tempdir", "s3://path/for/temp/data") \  
    .load()
```

## IAM を使用して認証情報を取得し、Amazon Redshift に接続する

Amazon Redshift が提供する JDBC バージョン 2 ドライバーを使用し、Spark コネクタで Amazon Redshift に接続できます。AWS Identity and Access Management (IAM) を使用するには、[IAM 認証を使用するように JDBC URL を設定します](#)。Amazon EMR から Redshift クラスターに接続するには、IAM ロールに一時的な IAM 認証情報を取得する権限を付与する必要があります。次の権限を IAM ロールに割り当てて、認証情報を取得して Amazon S3 オペレーションを実行できるようにします。

- [Redshift:GetClusterCredentials](#) (プロビジョニングされた Amazon Redshift クラスターの場合)
- [Redshift:DescribeClusters](#) (プロビジョニングされた Amazon Redshift クラスターの場合)
- [Redshift:GetWorkgroup](#) (Amazon Redshift Serverless ワークグループ用)
- [Redshift:GetCredentials](#) (Amazon Redshift Serverless ワークグループ用)
- [s3:GetBucket](#)
- [s3:GetBucketLocation](#)
- [s3:GetObject](#)
- [s3:PutObject](#)
- [s3:GetBucketLifecycleConfiguration](#)

`GetClusterCredentials` の詳細については、「[GetClusterCredentials のリソースポリシー](#)」を参照してください。

また、COPY および UNLOAD オペレーション中に Amazon Redshift が IAM ロールを引き継げるようにする必要があります。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "redshift.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

以下の例では、Spark と Amazon Redshift の間で IAM 認証を使用しています。

```
from pyspark.sql import SQLContext
import boto3

sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshift-host:redshift-port/db-name"
iam_role_arn = "arn:aws:iam::account-id:role/role-name"

# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("aws_iam_role", iam_role_arn) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3a://path/for/temp/data") \
    .mode("error") \
    .load()
```

## Amazon Redshift に対する読み書き

次のコード例では PySpark、を使用して、データソース API と SparkSQL を使用して Amazon Redshift データベースとの間でサンプルデータの読み取りと書き込みを行います。

### Data source API

PySpark データソース API を使用して、Amazon Redshift データベースとの間でサンプルデータを読み書きするには、を使用します。

```
import boto3
from pyspark.sql import SQLContext

sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshifthost:5439/database"
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
```

```
.option("url", url) \  
.option("dbtable", "tableName") \  
.option("tempdir", "s3://path/for/temp/data") \  
.option("aws_iam_role", "aws_iam_role_arn") \  
.load()  
  
df.write \  
.format("io.github.spark_redshift_community.spark.redshift") \  
.option("url", url) \  
.option("dbtable", "tableName_copy") \  
.option("tempdir", "s3://path/for/temp/data") \  
.option("aws_iam_role", "aws_iam_role_arn") \  
.mode("error") \  
.save()
```

## SparkSQL

PySpark を使用して、SparkSQL を使用して Amazon Redshift データベースとの間でサンプルデータを読み書きします。

```
import boto3  
import json  
import sys  
import os  
from pyspark.sql import SparkSession  
  
spark = SparkSession \  
.builder \  
.enableHiveSupport() \  
.getOrCreate()  
  
url = "jdbc:redshift:iam://redshifthost:5439/database"  
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"  
  
bucket = "s3://path/for/temp/data"  
tableName = "tableName" # Redshift table name  
  
s = f"""CREATE TABLE IF NOT EXISTS {tableName} (country string, data string)  
USING io.github.spark_redshift_community.spark.redshift  
OPTIONS (dbtable '{tableName}', tempdir '{bucket}', url '{url}', aws_iam_role  
'{aws_iam_role_arn}' ); """  
  
spark.sql(s)
```

```
columns = ["country" ,"data"]
data = [("test-country","test-data")]
df = spark.sparkContext.parallelize(data).toDF(columns)

# Insert data into table
df.write.insertInto(tableName, overwrite=False)
df = spark.sql(f"SELECT * FROM {tableName}")
df.show()
```

## Spark コネクタを使用する際の考慮事項と制限事項

- Amazon EMR 上の Spark から Amazon Redshift への JDBC 接続に対して SSL を有効にすることをお勧めします。
- ベストプラクティスとして、AWS Secrets Manager で Amazon Redshift クラスターの認証情報を管理することをお勧めします。例については[AWS Secrets Manager](#)、「[を使用して Amazon Redshift に接続するための認証情報を取得する](#)」を参照してください。
- Amazon Redshift 認証パラメータのパラメータ `aws_iam_role` を使用して IAM ロールを渡すことをお勧めします。
- 現在、パラメータ `tempformat` は Parquet 形式をサポートしていません。
- `tempdir` URI は Amazon S3 の場所を指します。この一時ディレクトリは、自動的にクリーンアップされないため、追加コストが発生する可能性があります。
- Amazon Redshift については、次の推奨事項を検討してください。
  - Amazon Redshift クラスターにパブリックにアクセスできないようにすることをお勧めします。
  - [Amazon Redshift 監査ログ作成](#) を有効にすることをお勧めします。
  - [Amazon Redshift 保管時の暗号化](#) を有効にすることをお勧めします。
- Amazon S3 については、次の推奨事項を検討してください。
  - [Amazon S3 バケットへのパブリックアクセスをブロックする](#) ことをお勧めします。
  - [Amazon S3 サーバー側の暗号化](#) を使用して、使用する Amazon S3 バケットを暗号化することをお勧めします。
  - [Amazon S3 ライフサイクルポリシー](#) を使用して、Amazon S3 バケットの保持ルールを定義することをお勧めします。
  - Amazon EMR は、常にオープンソースからイメージにインポートされるコードを検証します。セキュリティのため、Spark から Amazon S3 への次の認証方法はサポートされていません。

- `hadoop-env` 設定分類での AWS アクセスキーの設定
- `URI tempdir` での AWS アクセスキーのエンコード

コネクタとそのサポートされているパラメータの使用法の詳細については、次のリソースを参照してください。

- 「Amazon Redshift 管理ガイド」の「[Amazon Redshift integration for Apache Spark](#)」
- Github の [spark-redshift コミュニティリポジトリ](#)

## Spark リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Spark のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Important

Apache Spark バージョン 2.3.1 は Amazon EMR リリース 5.16.0 以降で利用でき、[CVE-2018-8024](#) および [CVE-2018-1334](#) に対処します。Spark の以前のバージョンを Spark バージョン 2.3.1 以降に移行することをお勧めします。

### Spark バージョン情報

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.36.2	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
		nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.1.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-7.0.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.15.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave



Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.14.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.13.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.12.0	3.4.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.11.1	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.11.0	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.10.1	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.10.0	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.9.1	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.9.0	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.8.1	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.8.0	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.7.0	3.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.36.1	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.36.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.6.0	3.2.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.35.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave



Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.5.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.4.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.3.1	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.3.0	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.2.1	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.2.0	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.1.1	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.1.0	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-6.0.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.0.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.34.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.33.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.33.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.32.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.32.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.31.1	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave



Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.31.0	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.2	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.30.1	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.0	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.29.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.28.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.28.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.27.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.27.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.26.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.25.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.24.1	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.24.0	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.23.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.23.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.22.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave



Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.21.2	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.21.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.21.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.20.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.20.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.19.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.19.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.18.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.18.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.2	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.17.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.16.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.16.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.15.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.15.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave



Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.14.2	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.14.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.14.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.13.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.13.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.12.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.12.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.4	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.11.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.11.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.10.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.10.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave



Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.9.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.9.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.8.3	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.2	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.8.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.7.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.7.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.6.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.6.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ftpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.5.4	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.3	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.5.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.4.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.4.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave



Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.3.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.3	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.2	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.2.1	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.0	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.1.1	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.1.0	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.3	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.2	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-5.0.1	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.0	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.6	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-4.9.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-4.9.2	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.1	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-4.8.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-4.8.1	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.0	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.4	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave



Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-4.7.3	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-4.7.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-4.5.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resource manager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.4.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resource manager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.3.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resource manager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR リリースラベル	Spark バージョン	Spark でインストールされるコンポーネント
emr-4.2.0	1.5.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.1.0	1.5.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.0.0	1.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-https-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

# Apache Sqoop

Apache Sqoop は、Amazon S3、Hadoop、HDFS、および RDBMS データベース間でデータを転送するためのツールです。詳細については、[Apache Sqoop](#) ウェブサイトを参照してください。Sqoop は、Amazon EMR リリース 5.0.0 以降に含まれています。以前のリリースには、サンドボックスアプリケーションとして Sqoop が含まれています。詳細については、「[Amazon EMR 4.x リリースバージョン](#)」を参照してください。

## トピック

- [Sqoop バージョン情報](#)
- [Amazon EMR での Sqoop に関する考慮事項](#)
- [Sqoop リリース履歴](#)

## Sqoop バージョン情報

### 7.1.0 の Sqoop バージョン

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Sqoop のバージョンと、Amazon EMR で Sqoop と共にインストールされるコンポーネントを示しています。

このリリースで Sqoop と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

### emr-7.1.0 の Sqoop バージョン情報

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-7.1.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
		timeline-server, mariadb-server, sqoop-client

### 6.15.0 用 Sqoop バージョン

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Sqoop のバージョンと、Amazon EMR で Sqoop と共にインストールされるコンポーネントを示しています。

このリリースで Sqoop と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Sqoop バージョン情報

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.15.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

### 5.36.2 の Sqoop バージョン

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Sqoop のバージョンと、Amazon EMR で Sqoop と共にインストールされるコンポーネントを示しています。

このリリースで Sqoop と共にインストールされるコンポーネントのバージョンについては、「[リリース 5.36.2 コンポーネントバージョン](#)」を参照してください。

## emr-5.36.2 の Sqoop バージョン情報

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.36.2	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

## Amazon EMR での Sqoop に関する考慮事項

Amazon EMR で Sqoop を実行するときは、以下について検討します。

### HCatalog の統合での Sqoop の使用

Amazon EMR の Sqoop は、[Sqoop-HCatalog の統合](#)をサポートします。Sqoop を使用して Amazon S3 の HCatalog テーブルに出力を書き込む場合は、`mapred.output.direct.NativeS3FileSystem` プロパティと `mapred.output.direct.EmrFileSystem` プロパティを `false` に設定して、Amazon EMR の直接書き込み機能を無効にしてください。詳細については、「[HCatalog の使用](#)」を参照してください。Hadoop の `-D mapred.output.direct.NativeS3FileSystem=false` および `-D mapred.output.direct.EmrFileSystem=false` コマンドを使用できます。直接書き込みを無効にしない場合、エラーは発生しませんが、テーブルが Amazon S3 で作成され、データは書き込まれません。

### Sqoop JDBC およびデータベースサポート

デフォルトでは、Sqoop には MariaDB および PostgreSQL ドライバーがインストールされています。Sqoop にインストールされた PostgreSQL ドライバーは PostgreSQL 8.4 に対してのみ機能します。Sqoop に代替となる一連の JDBC コネクタをインストールするには、クラスターのマスター

ノードに接続して `/usr/lib/sqoop/lib` にインストールします。以下は各種 JDBC コネクタのリンクです。

- MariaDB: [About MariaDB Connector/J](#)
- PostgreSQL: [PostgreSQL JDBC driver](#)。
- SQLServer: [Download Microsoft JDBC driver for SQL Server](#)。
- MySQL: [Download Connector/J](#)
- Oracle: [Get Oracle JDBC drivers and UCP from the Oracle Maven repository](#)

Sqoop でサポートされているデータベースは、[http://sqoop.apache.org/docs/\*version\*/.html#\\_supported\\_databases](http://sqoop.apache.org/docs/version/.html#_supported_databases) に一覧表示されます SqoopUserGuide。ここで、*version* は、使用している Sqoop のバージョンです。例えば、1.4.6 です。JDBC 接続文字列がこのリストの文字列に一致しない場合は、ドライバーを指定する必要があります。

例えば、次のコマンドを使用して Amazon Redshift データベーステーブルにエクスポートできます (JDBC 4.1 の場合)。

```
sqoop export --connect jdbc:redshift://$MYREDSHIFTHOST:5439/mydb --table mysqoopexport
--export-dir s3://mybucket/myinputfiles/ --driver com.amazon.redshift.jdbc41.Driver --
username master --password Mymasterpass1
```

MariaDB と MySQL 両方の接続文字列を使用できますが、MariaDB 接続文字列を指定する場合は、ドライバーを指定する必要があります。

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport
--export-dir s3://mybucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

Secure Socket Layer 暗号化を使用してデータベースに接続している場合、次の Sqoop エクスポート例にあるような JDBC URI を使用する必要があります。

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb?
verifyServerCertificate=false&useSSL=true&requireSSL=true --table mysqoopexport
--export-dir s3://mybucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

RDS での SSL 暗号化の詳細については、「Amazon RDS ユーザーガイド」の「[SSL を使用した DB インスタンスへの接続の暗号化](#)」を参照してください。



詳細については、[Apache Sqoop](#) ドキュメントを参照してください。

## パスワードの保護

パスワードを安全に渡す方法はいくつかあります。

### Java KeyStore

推奨される方法では、パスワードを Java KeyStore (JKS) で暗号化するため、パスワードを読み取り可能な形式で保存する必要はありません。

1. パスワードエイリアスを作成します。プロンプトが表示されたら、データベースへのアクセスに使用するパスワードを入力します。

```
hadoop credential create mydb.password.alias -provider jceks://hdfs/user/root/
mysql.password.jceks
```

2. パスワードエイリアスを使用して Sqoop ジョブを起動します。

```
sqoop export -Dhadoop.security.credential.provider.path=jceks://hdfs/user/
root/mysql.password.jceks --connect jdbc:mariadb://$HOSTNAME:3306/mydb
--table mysqoopexport --export-dir s3://mybucket/myinputfiles/ --driver
org.mariadb.jdbc.Driver --username master --password-alias mydb.password.alias
```

### --password-file

次の例に示すように、--password-file コマンドを使用してパスワードを渡すことができます。

1. パスワードを含む新しいファイルを作成します。

```
echo -n 'Mymasterpass1' > /home/hadoop/mysql-pass.password
```

2. このファイルを使用して Sqoop ジョブを起動します。

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport
--export-dir s3://mybucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password-file /home/hadoop/mysql-pass.password
```

-P

次の例に示すように、-P コマンドを使用して、プロンプトからパスワードを入力できます。

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://mybucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --username
master -P
```

## Sqoop リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Sqoop のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Sqoop バージョン情報

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.36.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
		httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.0.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.13.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.12.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.11.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.11.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.10.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.10.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.9.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.9.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.8.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client



Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.8.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.7.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.36.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.36.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.6.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.35.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.5.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.3.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.2.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-6.1.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.34.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.33.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client



Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.33.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.32.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.32.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.31.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.31.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.30.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.29.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.28.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.28.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.27.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.27.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.26.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.25.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.24.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client



Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.24.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.23.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.23.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.22.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.21.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.21.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.21.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.20.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.20.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.17.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client



Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.16.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.16.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.15.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.14.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.14.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.13.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.13.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.12.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.12.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.11.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client



Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.11.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.10.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.10.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.9.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.5.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.5.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client



Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.3.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.2.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.2.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされるコンポーネント
emr-5.0.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR リリースラベル	Sqoop バージョン	Sqoop でインストールされる コンポーネント
emr-5.0.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, mysql-server, sqoop-client

# TensorFlow

TensorFlow は、機械学習および深層学習アプリケーション用のオープンソースのシンボリック数学ライブラリです。詳細については、[TensorFlow ウェブサイト](#) TensorFlow 「」を参照してください。Amazon EMR リリースバージョン 5.17.0 以降で利用できます。

次の表に、Amazon EMR 7.x シリーズの最新リリース TensorFlow に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します TensorFlow。

このリリース TensorFlow で と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## TensorFlow emr-7.1.0 のバージョン情報

Amazon EMR リリースラベル	TensorFlow バージョン	でインストールされるコンポーネント TensorFlow
emr-7.1.0	TensorFlow 2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

次の表に、Amazon EMR 6.x シリーズの最新リリース TensorFlow に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します TensorFlow。

このリリース TensorFlow で と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## TensorFlow emr-6.15.0 のバージョン情報

Amazon EMR リリースラベル	TensorFlow バージョン	でインストールされるコンポーネント TensorFlow
emr-6.15.0	TensorFlow 2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

次の表に、Amazon EMR 5.x シリーズの最新リリース TensorFlow に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します TensorFlow。

このリリース TensorFlow で と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

## TensorFlow emr-5.36.2 のバージョン情報

Amazon EMR リリースラベル	TensorFlow バージョン	でインストールされるコンポーネント TensorFlow
emr-5.36.2	TensorFlow 2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow



## TensorFlow Amazon EC2 インスタンスタイプ別のビルド

Amazon EMR は、クラスター用に選択したインスタンスタイプに応じて、TensorFlow ライブラリのさまざまなビルドを使用します。次の表にインスタンスタイプ別のビルドを示します。

EC2 インスタンスタイプ	TensorFlow ビルド
M5 および C5	TensorFlow 1.9.0 (Intel MKL で最適化)
P2	TensorFlow 1.9.0 (CUDA 9.2、cuDNN 7.1 を使用)
P3	TensorFlow 1.9.0 (CUDA 9.2、cuDNN 7.1、NCCL 2.2.13 を使用)  <a href="#">NVIDIA NCCL</a> は、P3 インスタンスでのみ使用できます。エンドユーザーライセンス契約 (EULA): Amazon EMR で Nvidia のコンポーネントを使用すると、 <a href="#">製品の EULA</a> に概説されている諸条件に同意したことになります。
その他すべて	TensorFlow 1.9.0

## セキュリティ

「[TensorFlow を安全に使用する](#)」のガイダンスに従うことに加えて、信頼されたソースへのアクセスを制限できるように、プライベートサブネットでクラスターを起動することをお勧めします。詳細については、「Amazon EMR 管理ガイド」の「[Amazon VPC オプション](#)」を参照してください。

## の使用 TensorBoard

TensorBoard は、TensorFlow プログラム用の視覚化ツールのスイートです。詳細については、Tensorflow ウェブサイトの[TensorBoard 「: 視覚化学習」](#)を参照してください。

Amazon EMR TensorBoard でを使用するには、クラスターマスターノード TensorBoard でを起動する必要があります。

## Amazon EMR で TensorBoard と TensorFlow を使用するには

1. SSH を使用してクラスターのマスターノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してマスターノードに接続する](#)」を参照してください。
2. マスターノードで TensorBoard を起動するには、次のコマンドを入力します。サマリーライターを使用して概略データの生成と格納を行ったマスターノードのディレクトリで `/my/log/directory` を置き換えます。

### Amazon EMR 5.19.0 and later

```
python3 -m tensorboard.main --logdir=/home/hadoop/tensor --bind_all
```

### Amazon EMR 5.18.1 and earlier

```
python3 -m tensorboard.main --logdir=/my/log/dir
```

デフォルトでは、マスターノードはポート 6006 とマスターパブリック DNS 名 TensorBoard を使用してホストします。を起動すると TensorBoard、次の例に示すように TensorBoard、コマンドライン出力に への接続に使用できる URL が表示されます。

```
TensorBoard 1.9.0 at http://master-public-dns-name:6006 (Press CTRL+C to quit)
```

3. 信頼できるクライアントからマスターノードのウェブインターフェイスへのアクセスを設定します。詳細については、「Amazon EMR 管理ガイド」の「[Amazon EMR クラスターでホストされているウェブインターフェイスの表示](#)」を参照してください。
4. TensorBoard で を開きます `http://master-public-dns-name:6006`。

## TensorFlow リリース履歴

次の表に、Amazon EMR の各リリースバージョン TensorFlow に含まれる のバージョンと、アプリケーションと共にインストールされるコンポーネントを示します。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

## TensorFlow バージョン情報

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.36.2	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.1.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.0.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.15.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.14.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.13.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-6.12.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, tensorflow
emr-6.11.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, tensorflow
emr-6.11.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-6.10.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.10.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.1	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-6.9.0	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.8.1	2.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.8.0	2.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-6.7.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow



Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-6.6.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.35.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.5.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-6.4.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.3.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.3.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-6.2.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.2.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-6.1.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.0.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.0.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.34.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.33.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.33.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.32.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.32.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.31.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.2	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.30.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.29.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow



Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.28.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, tensorflow
emr-5.27.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, tensorflow
emr-5.27.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.26.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.25.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.24.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.23.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.23.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.22.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.2	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.21.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.20.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.20.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.19.1	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.19.0	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.18.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.2	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR リリースラベル	TensorFlow バージョン	と共にインストールされるコンポーネント TensorFlow
emr-5.17.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow



# Apache Tez

Apache Tez は、データを処理するためのタスクの複合 Directed Acyclic Graph (DAG) を作成するフレームワークです。一部の MapReduce ユースケースでは、Hadoop の代替として使用できます。例えば、Hadoop で Pig および Hive ワークフローを実行 MapReduce したり、Tez を実行エンジンとして使用したりできます。詳細については、<https://tez.apache.org/> を参照してください。Amazon EMR リリース 4.7.0 以降には Tez が含まれています。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Tez のバージョンと、Amazon EMR で Tez と共にインストールされるコンポーネントを示しています。

このリリースで Tez と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Tez バージョン情報

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-7.1.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Tez のバージョンと、Amazon EMR で Tez と共にインストールされるコンポーネントを示しています。

このリリースで Tez と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

## emr-6.15.0 の Tez バージョン情報

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-6.15.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Tez のバージョンと、Amazon EMR で Tez と共にインストールされるコンポーネントを示しています。

このリリースで Tez と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

## emr-5.36.2 の Tez バージョン情報

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.36.2	Tez 0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

## トピック

- [Tez を使用したクラスターの作成](#)
- [Tez の設定](#)
- [Tez ウェブ UI](#)
- [タイムラインサーバー](#)
- [Tez リリース履歴](#)

## Tez を使用したクラスターの作成

Tez をインストールするには、クラスターの作成時にアプリケーションとして Apache Tez を選択します。

コンソールを使用し、Tez がインストールされたクラスターを作成するには

1. 新しい Amazon EMR コンソールに移動し、サイドナビゲーションから [古いコンソールに切り替え] を選択します。古いコンソールに切り替えたときの動作の詳細については、「[Using the old console](#)」を参照してください。
2. [Create cluster (クラスターの作成)]、[Go to advanced options (詳細オプションに移動する)] の順に選択します。
3. [ソフトウェア設定] で、[emr-4.7.0] 以降の [リリース] を選択します。
4. [Tez] と Amazon EMR でインストールする他のアプリケーションを選択します。
5. 必要に応じて他のオプションを選択し、[Create cluster (クラスターの作成)] を選択します。

を使用して Tez でクラスターを作成するには AWS CLI

- `create-cluster` コマンドと `-- applications` オプションを使用して [Tez] を指定します。次の例では、Tez がインストールされたクラスターを作成します。

### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --name "Cluster with Tez" --release-label emr-7.1.0 \  
--applications Name=Tez --ec2-attributes KeyName=myKey \  
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

## Tez の設定

Tez をカスタマイズするには、値を設定して、tez-site 設定分類を使用します。これにより、tez-site.xml 設定ファイルが設定されます。詳細については、Apache Tez ドキュメント [TezConfiguration](#) の「」を参照してください。Hive または Pig を変更して Tez 実行エンジンを使用するには、必要に応じて hive-site と pig-properties 設定分類を使用します。次に例を示します。

### 設定例

Example 例: Tez のルートログインレベルを変更して Tez を Hive や Pig の実行エンジンとして設定する

次に示す create-cluster コマンドの例では、Tez、Hive、および Pig がインストールされたクラスターを作成します。このコマンドでは、Amazon S3 に格納されているファイル myConfig.json を参照します。そのファイルは tez-site 分類のプロパティを指定して tez.am.log.level を DEBUG に設定し、hive-site と pig-properties 設定分類を使用して Hive や Pig の実行エンジンを Tez に設定します。

#### Note

読みやすくするために、Linux 行連続文字 (\) が含まれています。Linux コマンドでは、これらは削除することも、使用することもできます。Windows の場合、削除するか、キャレット (^) に置き換えてください。

```
aws emr create-cluster --release-label emr-7.1.0 \  
--applications Name=Tez Name=Hive Name=Pig --ec2-attributes KeyName=myKey \  
--instance-type m5.xlarge --instance-count 3 \  
--configurations https://s3.amazonaws.com/mybucket/myfolder/myConfig.json --use-  
default-roles
```

myConfig.json のコンテンツの例を以下に示します。

```
[
  {
    "Classification": "tez-site",
    "Properties": {
      "tez.am.log.level": "DEBUG"
    }
  },
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.execution.engine": "tez"
    }
  },
  {
    "Classification": "pig-properties",
    "Properties": {
      "exectype": "tez"
    }
  }
]
```

#### Note

Amazon EMR バージョン 5.21.0 以降では、実行中のクラスター内のインスタンスグループごとに、クラスター設定を上書きして追加の設定分類を指定できます。これを行うには、Amazon EMR コンソール、AWS Command Line Interface ( AWS CLI )、または AWS SDK を使用します。詳細については、「[実行中のクラスター内のインスタンスグループの設定を指定する](#)」を参照してください。

## Tez 非同期分割オープニング

テーブルパスに小さなファイルが多数あり、クエリがそれらをすべて読み込もうとすると、個々の分割に対応する、それぞれの小さなファイルが 1 つの Tez グループ分割にまとめられます。次に、単一のマッパーが単一の Tez グループ化分割を処理します。実行は同期的に行われるため、グループ分割の下にある個々の分割は 1 つずつ処理されます。そのためには、RecordReader オブジェクトは分割を同期的に処理する必要があります。

名前	分類	説明
<code>tez.grouping.split.init.threads</code>	tez-site	Tez が RecordReaders およびオープン分割を事前に開始するために使用するデモンスレッドの数を指定します。ACID テーブルの場合、サポートされる <code>tez.grouping.split.init.threads</code> の最大値は 1 です。
<code>tez.grouping.split.init.recordreaders</code>	tez-site	デモンスレッドによって事前に初期化されたままにする RecordReaders の数を指定します。これは Tez のグループ分割に InputSplits が多数含まれている場合に役立ちます。このような入力分割を処理するための RecordReaders の初期化は、逐次処理ではなくデモンスレッドで非同期的に行うことができます。

## Tez 非同期分割オープンのベンチマーク

Tez の非同期分割オープン機能のベンチマークには、以下の環境と構成を使用しました。

- ベンチマーク環境 - m5.16xlarge を使用するプライマリノードが 1 つ、m5.16xlarge を使用するコノードが 16 個の Amazon EMR クラスター。
- ベンチマーク設定 - 単一の Tez グループ分割に多数の入力分割がある場合のベンチマークシナリオをシミュレートするには、`tez.grouping.split-count` を 1 に設定します。
- ベンチマークに使用されるテーブル - テーブルには 200 個のパーティションがあり、各パーティションには 1 つのファイルが含まれます。ベンチマークは、そのテーブルに CSV ファイルが含まれている場合と、そのテーブルにパーケットファイルが含まれている場合について行われます。ベ

ンチマーク用の Hive クエリ: テーブルから SELECT COUNT(\*) を 10 回実行し、平均実行時間を取得します。

- Tez 非同期分割オープンを有効にする設定 - 以下のとおりです。
  - `tez.grouping.split.init.threads = 4`
  - `tez.grouping.split.init.recordreaders = 10`

データセット	機能が無効 (ベースライン)	機能が有効	改良点
CSV データセット	90.26 秒	79.20 秒	12.25%
Parquet データセット	54.67 秒	42.23 秒	22.75%

## Tez ウェブ UI

Tez には、独自のウェブユーザーインターフェイスがあります。ウェブ UI を表示するには、次の URL を参照してください。

```
http://masterDNS:8080/tez-ui
```

Tez ウェブ UI で [Hive クエリ] タブを有効にするには、次の設定を指定します。

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.exec.pre.hooks": "org.apache.hadoop.hive.q1.hooks.ATSHook",
      "hive.exec.post.hooks": "org.apache.hadoop.hive.q1.hooks.ATSHook",
      "hive.exec.failure.hooks": "org.apache.hadoop.hive.q1.hooks.ATSHook"
    }
  }
]
```

コンソールでクラスターの詳細ページの [Application user interfaces] (アプリケーションユーザーインターフェイス) タブを使用して、Tez、Spark、YARN アプリケーション UI の詳細を表示することもできます。Amazon EMR アプリケーションユーザーインターフェイス (UI) はクラスター外でホス

トされ、クラスターが終了した後に使用できます。SSH 接続やウェブプロキシを設定する必要がなく、アクティブなジョブとジョブ履歴のトラブルシューティングおよび分析が容易になります。

詳細については、「Amazon EMR 管理ガイド」の「[アプリケーションの履歴を表示する](#)」を参照してください。

## タイムラインサーバー

YARN Timeline Server は、Tez がインストールされている場合に実行されるよう設定されます。Timeline Server を使用して Tez または MapReduce 実行エンジンを介して送信されたジョブを表示するには、URL を使用してウェブ UI を表示します `http://master-public-DNS:8188`。詳細については、「Amazon EMR 管理ガイド」の「[Amazon EMR クラスターでホストされているウェブインターフェイスの表示](#)」を参照してください。

## Tez リリース履歴

次の表は、Amazon EMR の各リリースに含まれている Tez のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Tez バージョン情報

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.36.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-7.1.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-



Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
		hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.0.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.15.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-6.14.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.13.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.12.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-6.11.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.11.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-6.10.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.9.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.9.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-6.8.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.8.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.7.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.36.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.36.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.6.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.35.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.5.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.4.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-6.3.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.3.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn



Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-6.2.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.1.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.1.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-6.0.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.0.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.34.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.33.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.33.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.32.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.31.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.31.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.30.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.29.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.28.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.28.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.27.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.27.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.26.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.25.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.24.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.24.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn



Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.23.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.23.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.22.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.21.2	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.20.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.20.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.19.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.18.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.18.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.17.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.16.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.16.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.15.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.14.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.13.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.13.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn



Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.12.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.12.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.11.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.10.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.10.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.9.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.7.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.6.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.6.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.5.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn



Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.5.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.4.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.3.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.2.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.1.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-5.0.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.6	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-4.9.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-4.9.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn



Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-4.8.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-4.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.4	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-4.7.3	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.2	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.1	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR リリースラベル	Tez バージョン	Tez でインストールされるコンポーネント
emr-4.7.0	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resource-manager, hadoop-yarn-timeline-server, tez-on-yarn

## Tez リリースノート (バージョン別)

### トピック

- [Amazon EMR 7.1.0 - Tez リリースノート](#)
- [Amazon EMR 6.15.0 - Tez リリースノート](#)
- [Amazon EMR 6.14.0 - Tez リリースノート](#)
- [Amazon EMR 6.13.0 - Tez リリースノート](#)
- [Amazon EMR 6.12.0 - Tez リリースノート](#)
- [Amazon EMR 6.11.0 - Tez リリースノート](#)
- [Amazon EMR 6.10.0 - Tez リリースノート](#)
- [Amazon EMR 6.9.0 - Tez リリースノート](#)
- [Amazon EMR 6.8.0 - Tez リリースノート](#)
- [Amazon EMR 6.7.0 - Tez リリースノート](#)
- [Amazon EMR 6.6.0 - Tez リリースノート](#)

## Amazon EMR 7.1.0 - Tez リリースノート

### Amazon EMR 7.1.0 - Tez の変更点

型	説明
バグ修正	<a href="#">TEZ-4394</a> – Netty4 ShuffleHandler: は 1 つのボススレッド (#256) を使用する必要があります
Upgrade	Netty を 4.1.100.Final にアップグレードします。
Upgrade	Jetty を 9.4.53.v20231009 にアップグレードします。

## Amazon EMR 6.15.0 - Tez リリースノート

### Amazon EMR 6.15.0 - Tez の変更点

型	説明
機能	<a href="#">TEZ-4397</a> : オープン Tez 入力が非同期的に分割される
Upgrade	<a href="#">TEZ-4493</a> : Apache Hadoop を 3.3.6 にアップグレード

### Amazon EMR 6.15.0 - Tez 機能

- [Tez 非同期分割オープン](#) - Amazon EMR 6.15.0 では、入力分割を Tez グループ分割で非同期的に開くように指定できる構成が導入されています。この機能は [TEZ-4397](#) で開始されましたが、OSS Hive ではリグレッションが発生していました。Amazon EMR Hive では Hive ACID テーブルのリグレッションとその他のバグが修正されました。この改善により、1 つの Tez グループ分割で多数の入力分割がある場合に、読み取りクエリのパフォーマンスが向上します。詳細については、「[Tez 非同期分割オープニング](#)」を参照してください。

## Amazon EMR 6.14.0 - Tez リリースノート

### Amazon EMR 6.14.0 - Tez の変更点

型	説明
改良点	Tez の TLS バージョンを 1.2 に更新します

## Amazon EMR 6.13.0 - Tez リリースノート

### Amazon EMR 6.13.0 - Tez の変更点

型	説明
バグ修正	REVERT <a href="#">TEZ-4295</a> : データを解凍できませんでした。バッファ長が小さすぎます。
バグ修正	REVERT <a href="#">TEZ-4302</a> : GzipCodec を伴う CodecUtils での NullPointerException 。
バグ修正	REVERT <a href="#">TEZ-4234</a> : 制限が容量を超えると、コンプレッサーが Buffer.limit での IllegalArgumentException の原因となることがあります。
バグ修正	REVERT <a href="#">TEZ-4135</a> : インメモリ読み取りの実行時のメモリ割り当てを改善します。

## Amazon EMR 6.12.0 - Tez リリースノート

### Amazon EMR 6.12.0 - Tez の変更点

型	説明
改良点	JDK 11 と JDK 17 ランタイムのサポートが追加されました

型	説明
バグ修正	<a href="#">TEZ-4492</a> : Bower レジストリ CERT_EXPIRE の問題 (BOWER-2608) を回避するため、bower.herokuapp ミラーを使用するように Bowerrc を更新します
Upgrade	Surefire を 3.0.0-M7 にアップグレードしました

## Amazon EMR 6.11.0 - Tez リリースノート

### Amazon EMR 6.11.0 - Tez の変更点

型	説明
バグ	シャッフルデータの頂点レベルクリーンアップ中の無効な頂点状態遷移が修正されました
バグ	シャッフルデータの DAG または頂点レベルのクリーンアップが機能しない問題が修正されました
改良点	完了した DAG のシャッフルデータをクリアするために <code>tez.am.dag.cleanup.on.completion</code> をデフォルトで有効にします

## Amazon EMR 6.10.0 - Tez リリースノート

### Amazon EMR 6.10.0 - Tez の変更点

型	説明
機能	デフォルトで <code>tez.runtime.transfer.data-via-events.enabled</code> を有効にします

型	説明
バックポート	<a href="#">TEZ-4450</a> : シャッフルデータがデータ移動イベント経由で転送された際のシャッフルデータ取得の失敗を修正します
バックポート	<a href="#">TEZ-4460</a> : Tez シャッフルハンドラーからシャッフルデータを取得する際の読み取りタイムアウトエラーを修正します
バックポート	<a href="#">TEZ-4455</a> : デバッグ性を高めるため LoggingHandler に ShuffleHandlerパイプラインに追加する
バグ	タスクのプリエンプションが有効になっていると、Tez タスクが断続的に停止してしまう問題を修正します

## Amazon EMR 6.9.0 - Tez リリースノート

### Amazon EMR 6.9.0 - Tez の変更点

型	説明
Upgrade	Tez は 0.10.2 にアップグレードされました。詳細については、「 <a href="#">change log for Apache Tez 0.10.2</a> 」を参照してください。
Upgrade	Hadoop を 3.3.3 に更新します。
バグ	<a href="#">TEZ-4450</a> のため、 <code>tez.runtime.transfer.data-via-events.enabled</code> をデフォルトで無効にします。



## Amazon EMR 6.8.0 - Tez リリースノート

### Amazon EMR 6.8.0 - Tez の変更点

型	説明
バックポート	<a href="#">TEZ-3363</a> : シャッフルハンドラーで頂点レベルの中間データを削除します
バックポート	<a href="#">TEZ-4129</a> : シャッフルハンドラーで失敗した試行の中間試行データを削除します
バックポート	<a href="#">TEZ-4430</a> : tez.task.launch.cmd-opts プロパティが動作しない問題が修正されました

## Amazon EMR 6.7.0 - Tez リリースノート

### Amazon EMR 6.7.0 - Tez の変更点

型	説明
バックポート	<a href="#">TEZ-4403</a> : SLF4J のバージョンを 1.7.36 に更新します
バックポート	<a href="#">TEZ-4405</a> : log4j 1.x を reload4j に置き換えます
バックポート	<a href="#">TEZ-4411</a> : Tez ビルド失敗 : FileSaver.js が見つかりません

## Amazon EMR 6.6.0 - Tez リリースノート

### Amazon EMR 6.6.0 - Tez の変更点

型	説明
バックポート	<a href="#">TEZ-3918</a> : tez.task.log.level プロパティが動作しない問題が修正されました。

型	説明
バックポート	<a href="#">TEZ-4353</a> : commons-io を 2.8.0 に更新します。
バックポート	<a href="#">TEZ-4114</a> : tez から jetty の直接依存を削除します。
バックポート	<a href="#">TEZ-4323</a> : TEZ-4114 で Jetty jars が dist パッケージから削除されました。

# Apache Zeppelin

インタラクティブなデータ検索用のノートブックとして Apache Zeppelin を使用します。Zeppelin の詳細については、<https://zeppelin.apache.org/> を参照してください。Zeppelin は、Amazon EMR リリース 5.0.0 以降に含まれています。以前のリリースには、サンドボックスアプリケーションとして Zeppelin が含まれています。詳細については、「[Amazon EMR 4.x リリースバージョン](#)」を参照してください。

Zeppelin ウェブインターフェイスにアクセスするには、マスターノードへの SSH トンネルとプロキシ接続を設定します。詳細については、「[EMR クラスタでホストされているウェブサイトの表示](#)」を参照してください。

次の表は、Amazon EMR 7.x シリーズの最新リリースに含まれている Zeppelin のバージョンと、Amazon EMR で Zeppelin と共にインストールされるコンポーネントを示しています。

このリリースで Zeppelin と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## emr-7.1.0 の Zeppelin バージョン情報

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-7.1.0	Zeppelin 0.10.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

次の表は、Amazon EMR 6.x シリーズの最新リリースに含まれている Zeppelin のバージョンと、Amazon EMR で Zeppelin と共にインストールされるコンポーネントを示しています。

このリリースで Zeppelin と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。

#### emr-6.15.0 の Zeppelin バージョン情報

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.15.0	Zeppelin 0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

次の表は、Amazon EMR 5.x シリーズの最新リリースに含まれている Zeppelin のバージョンと、Amazon EMR で Zeppelin と共にインストールされるコンポーネントを示しています。

このリリースで Zeppelin と共にインストールされるコンポーネントのバージョンについては、「[リリース 5.36.2 コンポーネントバージョン](#)」を参照してください。

#### emr-5.36.2 の Zeppelin バージョン情報

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.36.2	Zeppelin 0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
		client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

## トピック

- [Amazon EMR で Zeppelin を使用するときの考慮事項](#)
- [Zeppelin リリース履歴](#)

## Amazon EMR で Zeppelin を使用するときの考慮事項

- 同じ [SSH トンネリングメソッド](#) を使用して Zeppelin に接続し、マスターノードの他のウェブサーバーに接続します。Zeppelin サーバーはポート 8890 です。
- Amazon EMR リリースバージョン 5.0.0 以降の Zeppelin は、[Shiro 認証](#) をサポートしています。
- Amazon EMR リリースバージョン 5.8.0 以降の Zeppelin では、Spark SQL のメタストアとして AWS Glue Data Catalog の使用がサポートされています。詳細については、[「Spark SQL のメタストアとして AWS Glue Data Catalog を使用する」](#) を参照してください。
- spark.dynamicAllocation.enabled が true に設定されている場合、Zeppelin は YARN に対してエグゼキュターを動的に割り当てるように指示をしますが、クラスターの spark-defaults.conf 設定ファイルで定義された設定の一部は使用しません。Zeppelin の [Interpreter (インタープリタ)] タブを使用してメモリやコアなどのエグゼキュターの設定を行い、インタープリタを再起動してそれらを使用できるようにする必要があります。

- Amazon EMR リリース 6.10.0 以降では、Apache Zeppelin と Apache Flink の統合がサポートされています。詳細については、「[Amazon EMR の Zeppelin から Flink ジョブを操作する](#)」を参照してください。
- Amazon EMR 上の Zeppelin は SparkR インタープリタをサポートしていません。

## Zeppelin リリース履歴

次の表は、Amazon EMR の各リリースバージョンに含まれている Zeppelin のバージョンと、アプリケーションと共にインストールされるコンポーネントを示しています。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

### Zeppelin バージョン情報

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.36.2	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.1.0	0.10.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
		<p>httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server</p>
emr-7.0.0	0.10.1	<p>emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server</p>

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.15.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.14.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server



Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.13.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.12.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.11.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.11.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.10.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.10.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.9.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.9.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.8.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.8.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.7.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.36.1	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.36.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.6.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.35.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.5.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server



Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.4.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.3.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.3.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.2.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.2.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.1.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.1.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.0.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-6.0.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.34.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.33.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.33.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.32.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.32.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.31.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.31.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server



Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.30.2	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.30.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.30.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.29.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.28.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.28.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.27.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.27.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.26.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.25.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.24.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.24.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.23.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.23.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.22.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.2	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server



Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.21.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.20.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.20.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.19.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.19.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.18.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.18.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.17.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.17.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.17.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.16.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.16.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.15.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.15.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server



Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.14.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.13.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.13.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.12.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.12.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.11.4	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.11.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.11.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.10.1	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.10.0	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.9.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server



Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.9.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.3	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.8.2	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.8.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.7.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.7.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.6.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.6.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.4	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.5.3	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.2	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.5.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.4.1	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.4.0	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server



Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.3.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.3.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.3.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.3	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.2.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.2.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.1.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.1.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.0.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR リリースラベル	Zeppelin バージョン	Zeppelin と共にインストールされるコンポーネント
emr-5.0.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

# Apache ZooKeeper

Apache ZooKeeper は、設定情報の維持、命名、分散同期の提供、グループサービスの提供のための一元化されたサービスです。の詳細については ZooKeeper、<http://zookeeper.apache.org/> を参照してください。

次の表に、Amazon EMR 7.x シリーズの最新リリース ZooKeeper に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します ZooKeeper。

このリリース ZooKeeper で と共にインストールされるコンポーネントのバージョンについては、「[リリース 7.1.0 コンポーネントバージョン](#)」を参照してください。

## ZooKeeper emr-7.1.0 のバージョン情報

Amazon EMR リリースラベル	ZooKeeper バージョン	でインストールされるコンポーネント ZooKeeper
emr-7.1.0	ZooKeeper 3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

次の表に、Amazon EMR 6.x シリーズの最新リリース ZooKeeper に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します ZooKeeper。

このリリース ZooKeeper で と共にインストールされるコンポーネントのバージョンについては、「[リリース 6.15.0 コンポーネントバージョン](#)」を参照してください。



## ZooKeeper emr-6.15.0 のバージョン情報

Amazon EMR リリースラベル	ZooKeeper バージョン	でインストールされるコンポーネント ZooKeeper
emr-6.15.0	ZooKeeper 3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

次の表に、Amazon EMR 5.x シリーズの最新リリース ZooKeeper に含まれている のバージョンと、Amazon EMR が と共にインストールするコンポーネントを示します ZooKeeper。

このリリース ZooKeeper で と共にインストールされるコンポーネントのバージョンについては、[「リリース 5.36.2 コンポーネントバージョン」](#)を参照してください。

## ZooKeeper emr-5.36.2 のバージョン情報

Amazon EMR リリースラベル	ZooKeeper バージョン	でインストールされるコンポーネント ZooKeeper
emr-5.36.2	ZooKeeper 3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

## トピック

- [ZooKeeper リリース履歴](#)

## ZooKeeper リリース履歴

次の表に、Amazon EMR の各リリースバージョン ZooKeeper に含まれる のバージョンと、アプリケーションと共にインストールされるコンポーネントを示します。各リリースのコンポーネントバージョンについては、[Amazon EMR 7.x リリースバージョン](#)、[Amazon EMR 6.x リリースバージョン](#) または [Amazon EMR 5.x リリースバージョン](#) でリリースに応じたコンポーネントバージョンのセクションを参照してください。

## ZooKeeper バージョン情報

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.36.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.1.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-7.0.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.15.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.14.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-6.13.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.12.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.11.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-6.11.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.10.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.10.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-6.9.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.9.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.8.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-6.8.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.7.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.36.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.6.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.35.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-6.5.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.4.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.3.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-6.3.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.2.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.2.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-6.1.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.1.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.0.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-6.0.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.34.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.33.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.33.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.32.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.32.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.31.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.31.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.30.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.29.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.28.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.28.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.27.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.27.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.26.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.25.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.24.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.24.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.23.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.23.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.22.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.2	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.21.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.20.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.20.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.19.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.19.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.18.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.18.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.2	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.17.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.16.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.16.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.15.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.15.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.14.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.13.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.13.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.12.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.11.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.11.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.10.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.10.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.9.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.9.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.8.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.8.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.7.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.7.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server



Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.6.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.6.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.5.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.5.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.5.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.4.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.4.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.3.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.2.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.2.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.1.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.1.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.0.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-https-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR リリースラベル	ZooKeeper バージョン	と共にインストールされるコンポーネント ZooKeeper
emr-5.0.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server



# コネクタおよびユーティリティ

Amazon EMR には、データソースとして他の AWS のサービスにアクセスするためのコネクタとユーティリティがいくつか用意されています。これらのサービスのデータには、通常プログラム内でアクセスします。例えば、Hive クエリ、Pig スクリプト、または MapReduce アプリケーションで Kinesis ストリームを指定し、そのデータを操作できます。

## トピック

- [Amazon EMR による DynamoDB 内テーブルのエクスポート、インポート、クエリ、結合](#)
- [Kinesis](#)
- [S3DistCp \(s3-dist-cp\)](#)
- [S3DistCp ジョブが失敗した後のクリーンアップ](#)

## Amazon EMR による DynamoDB 内テーブルのエクスポート、インポート、クエリ、結合

### Note

Amazon EMR-DynamoDB コネクタは、[でオープンソース化されています](https://github.com/aws-labs/emr-dynamodb-connector) GitHub。詳細については、<https://github.com/aws-labs/emr-dynamodb-connector> を参照してください。

DynamoDB は、高速で予測可能なパフォーマンスとシームレスな拡張性を特長とするフルマネージド NoSQL データベースサービスです。開発者がデータベーステーブルを作成し、そのリクエストトラフィックやストレージを無制限に拡張できます。DynamoDB は、高速処理を維持しながら、テーブルのデータとトラフィックを十分な数のサーバーに自動的に分散し、顧客が指定したリクエスト容量と、格納されているデータ量を処理します。Amazon EMR および Hive を使用すると、DynamoDB に格納されているデータなど、大量のデータを迅速かつ効率的に処理できます。DynamoDB の詳細については、「[Amazon DynamoDB 開発者ガイド](#)」を参照してください。

Apache Hive は、HiveQL と呼ばれる簡素化された SQL のようなクエリ言語を使用して、マップリデュースクラスターにクエリを実行する際に使用できるソフトウェア層で、Hadoop アーキテクチャーの上で実行されます。Hive および HiveQL の詳細については、「[HiveQL 言語マニュアル](#)」を参照してください。Hive および Amazon EMR の詳細については、「[Apache Hive](#)」を参照してください。

DynamoDB への接続を含む、カスタマイズされたバージョンの Hive とともに Amazon EMR を使用して、DynamoDB に保存されたデータに対してオペレーションを実行できます。

- DynamoDB データを Hadoop Distributed File System (HDFS) にロードし、そのデータを Amazon EMR クラスターへの入力として使用する。
- SQL のようなステートメント ( HiveQL ) を使用してライブ DynamoDB データにクエリを実行する
- DynamoDB に格納されたデータを結合し、そのデータをエクスポートするか、結合されたデータに対してクエリを実行する。
- DynamoDB に格納されたデータを Amazon S3 にエクスポートする
- Amazon S3 に格納されたデータを DynamoDB にインポートする

#### Note

Amazon EMR-DynamoDB コネクタは、[Kerberos 認証](#)を使用するように設定されたクラスターをサポートしていません。

次のタスクのいずれかを実行するには、Amazon EMR クラスターを起動し、DynamoDB のデータの場所を指定して、DynamoDB のデータを操作するための Hive コマンドを発行します。

Amazon EMR クラスターを起動するには、Amazon EMR コンソール、コマンドラインインターフェイス (CLI)、または AWS SDK または Amazon EMR API を使用してクラスターをプログラムする方法がいくつかあります。また、Hive クラスターをインタラクティブに実行するか、スクリプトから実行するかを選択することもできます。このセクションでは、インタラクティブな Hive クラスターを Amazon EMR コンソールと CLI から起動する方法について説明します。

Hive をインタラクティブに使用方法は、クエリパフォーマンスをテストし、アプリケーションを調整するには最適です。定期的に行う一連の Hive コマンドを確立したら、Amazon EMR が実行できる Hive スクリプトを作成することを検討します。

#### Warning

DynamoDB テーブルに対する Amazon EMR 読み取りまたは書き込みオペレーションは、確立されたプロビジョニング済みスループットに対して不利に働き、プロビジョニングされたスループットの例外が発生する頻度が高くなる可能性があります。リクエストが多い場合、Amazon EMR はエクスポネンシャルバックオフによってリトライを実装

し、DynamoDB テーブルに対するリクエストロードを管理します。Amazon EMR ジョブを他のトラフィックと同時に実行すると、割り当てられたプロビジョニング済みスループットレベルを超えることがあります。これは、Amazon の ThrottleRequests メトリクスを確認することでモニタリングできます CloudWatch。リクエストロードが高すぎる場合は、クラスターを再作成して、[読み込みパーセントの設定](#) または [書き込みパーセントの設定](#) をより低い値に設定し、Amazon EMR オペレーションを調整できます。DynamoDB スループットの設定については、「[プロビジョニングされたスループット](#)」を参照してください。テーブルが [オンデマンドモード](#) に設定されている場合は、エクスポートまたはインポート操作を実行する前に、テーブルをプロビジョニングモードに戻す必要があります。DynamoDBTable から使用するリソースを計算するには、パイプラインにスループット比率が必要です。オンデマンドモードでは、プロビジョニングされたスループットが削除されます。スループットキャパシティをプロビジョニングするには、Amazon CloudWatch Events メトリクスを使用して、テーブルが使用した総スループットを評価できます。

## トピック

- [Hive コマンドを実行するように Hive テーブルをセットアップする](#)
- [DynamoDB 内データのエクスポート、インポート、クエリを行う Hive コマンドの使用例](#)
- [DynamoDB での Amazon EMR オペレーションのパフォーマンスの最適化](#)

## Hive コマンドを実行するように Hive テーブルをセットアップする

Apache Hive は、SQL のような言語を使用して、Amazon EMR クラスターに含まれるデータに対してクエリを実行する際に使用できるデータウェアハウスアプリケーションです。Hive の詳細については、<http://hive.apache.org/> を参照してください。

次の手順は、既にクラスターを作成し、Amazon EC2 キーペアを指定したことを前提にしています。クラスターの作成を開始する方法については、「Amazon EMR 管理ガイド」の「[Amazon EMR の開始方法](#)」を参照してください。

### を使用するように Hive を設定する MapReduce

Amazon EMR で Hive を使用して DynamoDB テーブルに対してクエリを実行する場合、デフォルトの実行エンジンである Tez を Hive が使用していると、エラーが発生する可能性があります。このため、このセクションで説明するように DynamoDB と統合する Hive でクラスターを作成する場合は、を使用するように Hive を設定する設定分類を使用することをお勧めします MapReduce。詳細については、「[アプリケーションの設定](#)」を参照してください。

次のスニペットは、を Hive の実行エンジン MapReduce として設定するために使用する設定分類とプロパティを示しています。

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.execution.engine": "mr"
    }
  }
]
```

Hive コマンドをインタラクティブに実行するには

1. マスターノードに接続します。詳細については、「Amazon EMR 管理ガイド」の「[SSH を使用してマスターノードに接続する](#)」を参照してください。
2. 現在のマスターノードの コマンドプロンプトで、hive と入力します。

hive> という Hive プロンプトが表示されます。

3. Hive アプリケーションのテーブルを DynamoDB のデータにマップする Hive コマンドを入力します。このテーブルは、Amazon DynamoDB に格納されているデータの参照として機能します。データは Hive のローカルに保存されておらず、このテーブルを使用するクエリは、DynamoDB 内のライブデータに対して実行されます。また、コマンドを実行するたびに、テーブルの読み取り容量と書き込み容量が消費されます。同じデータセットに対して複数の Hive コマンドを実行する予定がある場合は、まずエクスポートすることをお勧めします。

次に、Hive テーブルを DynamoDB テーブルにマッピングする構文を示します。

```
CREATE EXTERNAL TABLE hive_tablename
  (hive_column1_name column1_datatype, hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
"hive_column1_name:dynamodb_attribute1_name, hive_column2_name:dynamodb_attribute2_name...")
```

DynamoDB から Hive にテーブルを作成する場合、EXTERNAL というキーワードを使用して外部テーブルとして作成する必要があります。外部テーブルと内部テーブルの違いは、内部テーブ

ルの drop が実行されると内部テーブルのデータが削除される点です。Amazon DynamoDB に接続する場合、この動作は望ましくないため、外部テーブルのみがサポートされます。

たとえば、次の Hive コマンドでは、dynamodhtable1 という DynamoDB テーブルを参照する hivetable1 というテーブルが Hive に作成されます。DynamoDB テーブル dynamodhtable1 には hash-and-range プライマリキースキーマがあります。ハッシュキー要素は name ( 文字列型 )、範囲キー要素は year ( 数値型 ) です。各項目には、holidays ( 文字列セット型 ) の属性値があります。

```
CREATE EXTERNAL TABLE hivetable1 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodhtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");
```

行 1 では、HiveQL CREATE EXTERNAL TABLE ステートメントを使用しています。hivetable1 の場合、DynamoDB テーブルの属性名と値の各ペアについて、列を設定し、データ型を指定する必要があります。これらの値は大文字と小文字が区別されません。列には、( 予約語を除き ) 任意の名前を付けることができます。

行 2 では、STORED BY ステートメントを使用しています。STORED BY の値は、Hive と DynamoDB 間の接続を処理するクラス名です。'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler' のように設定する必要があります。

行 3 では、TBLPROPERTIES ステートメントを使用して、"hivetable1" を DynamoDB 内の正しいテーブルとスキーマに関連付けています。TBLPROPERTIES に、dynamodb.table.name パラメータと dynamodb.column.mapping パラメータの値を指定します。これらの値は、大文字と小文字が区別されます。

#### Note

テーブルのすべての DynamoDB 属性名には、Hive テーブルに対応する列が存在する必要があります。Amazon EMR のバージョンによっては、one-to-one マッピングが存在しない場合に次のシナリオが発生します。

- Amazon EMR バージョン 5.27.0 以降では、コネクタには、DynamoDB 属性名と Hive テーブルの列間の one-to-one マッピングを保証する検証があります。one-to-one マッピングが存在しない場合、エラーが発生します。
- Amazon EMR バージョン 5.26.0 以前では、Hive テーブルには DynamoDB からの名前と値のペアは含まれません。DynamoDB プライマリキー属性をマップしない場合、Hive からエラーが生成されます。プライマリキー以外の属性をマップしない場合、エラーは生成されませんが、Hive テーブルのデータは表示されません。データ型が一致しない場合、値は null です。

次に、hivetable1 で Hive 操作の実行を開始できます。hivetable1 に対して実行されるクエリは、DynamoDB アカウントの DynamoDB テーブル dynamoddbtable1 に対して内部的に実行され、実行するたびに読み取り単位または書き込み単位が消費されます。

DynamoDB テーブルに対して Hive クエリを実行する際には、十分な量の読み込み容量単位をプロビジョニングしておく必要があります。

たとえば、DynamoDB テーブルに対して 100 ユニットの読み込みキャパシティーをプロビジョニングしているとします。この場合、1 秒間に 100 の読み込み ( 409,600 バイト ) を実行できます。そのテーブルに 20 GB ( 21,474,836,480 バイト ) のデータが含まれており、Hive クエリがフルテーブルスキャンを実行する場合、クエリの実行にかかる時間は次のように見積もられます。

$$21,474,836,480 / 409,600 = 52,429 \text{ 秒} = 14.56 \text{ 時間}$$

必要な時間を短縮するには、ソース DynamoDB テーブルで読み込みキャパシティーユニットを調整する以外に方法はありません。Amazon EMR ノードを追加しても、役に立ちません。

Hive 出力では、1 つ以上のマッパープロセスが終了すると、完了のパーセンテージが更新されます。プロビジョニングされた読み込みキャパシティーが小さく設定された大きな DynamoDB テーブルでは、完了のパーセンテージ出力が長時間更新されない場合があります。上記のような場合、ジョブは数時間にわたって 0% 完了として表示されます。ジョブの進行状況の詳細なステータスについては、Amazon EMR コンソールに移動してください。ここで、個別のマッパータスクのステータスおよびデータ読み込みの統計を表示できます。また、マスターノードの Hadoop インターフェイスにログオンし、Hadoop 統計を表示することもできます。ここには、個別のマッパータスクのステータスおよびいくつかのデータ読み込み統計が表示されます。詳細については、次のトピックを参照してください。

- [マスターノード上にホストされるウェブインターフェイス](#)

## • [Hadoop のウェブインターフェイスを表示する](#)

DynamoDB からのデータのエクスポートやインポート、テーブルの結合などのタスクを実行するその他の HiveQL ステートメントの例については、「[DynamoDB 内データのエクスポート、インポート、クエリを行う Hive コマンドの使用例](#)」を参照してください。

Hive リクエストをキャンセルするには

Hive クエリを実行すると、サーバーから返される最初の応答には、リクエストをキャンセルするコマンドが含まれます。プロセスの任意の時点でリクエストをキャンセルするには、サーバーの応答に含まれる Kill Command を使用します。

1. Ctrl+C を入力して、コマンドラインクライアントを終了します。
2. シェルプロンプトで、リクエストに対するサーバーからの最初の応答に含まれていた Kill Command を入力します。

または、マスターノードのコマンドラインから次のコマンドを実行して、Hadoop ジョブをキルします。この *job-id* は、Hadoop ジョブの識別子であり、Hadoop ユーザーインターフェイスから取得できます。

```
hadoop job -kill job-id
```

## Hive と DynamoDB のデータ型

次の表は、使用可能な Hive データ型、対応するデフォルトの DynamoDB 型、およびそれらもマッピングできる代替 DynamoDB 型を示しています。

Hive の型	デフォルトの DynamoDB 型	代替 DynamoDB 型
string	文字列 ( S )	
bigint または double	数値 ( N )	
バイナリ	バイナリ ( B )	



Hive の型	デフォルトの DynamoDB 型	代替 DynamoDB 型
ブール値	boolean (BOOL)	
array	list (L)	数値セット ( NS )、文字列セット ( SS )、またはバイナリセット ( BS )
map<string,string>	item	map (M)
map<string,?>	map (M)	
	null (NULL)	

Hive データを対応する代替 DynamoDB 型として書き込む場合、または DynamoDB データに代替 DynamoDB 型の属性値が含まれる場合、`dynamodb.type.mapping` パラメータを使用して列と DynamoDB 型を指定できます。次の例は、代替型マッピングを指定する構文を示しています。

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
"hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.type.mapping" = "hive_column1_name:dynamodb_attribute1_datatype");
```

型マッピングパラメータはオプションで、代替型を使用する列に対してのみ指定する必要があります。

たとえば、次の Hive コマンドでは、DynamoDB テーブル `dynamodhtable2` を参照する `hivetable2` というテーブルが作成されます。これは、`col3` 列を文字列セット (SS) 型にマッピングする点を除いて、`hivetable1` に似ています。

```
CREATE EXTERNAL TABLE hivetable2 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodhtable2",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays",
"dynamodb.type.mapping" = "col3:SS");
```



Hive では、`hivetable1` と `hivetable2` は同一です。ただし、これらのテーブルからのデータが対応する DynamoDB テーブルに書き込まれる際、`dynamodatabase2` には文字列セットが含まれますが `dynamodatabase1` にはリストが含まれます。

Hive `null` 値を DynamoDB `null` 型の属性として書き込む場合は、`dynamodb.null.serialization` パラメータを使用して実行できます。次の例は、`null` シリアル化を指定する構文を示しています。

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,  
hive_column2_name column2_datatype...)  
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'  
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",  
"dynamodb.column.mapping" =  
"hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",  
"dynamodb.null.serialization" = "true");
```

`null` シリアル化パラメータはオプションであり、指定しない場合は `false` に設定されます。パラメータ設定に関係なく、DynamoDB `null` 属性は Hive で `null` 値として読み取られることに注意してください。`null` 値を持つ Hive コレクションは、`null` シリアル化パラメータが `true` として指定されている場合にのみ DynamoDB に書き込むことができます。指定されない場合、Hive エラーが発生します。

精度の点で、Hive の `bigint` 型は Java の `long` 型と同じであり、Hive の `double` 型は Java の `double` 型と同じです。つまり、エクスポート、インポート、または参照する Hive を使用して、Hive のデータ型で利用できる精度よりも高い精度の数値データが DynamoDB に格納される場合、DynamoDB データの精度が低下したり、Hive クエリが失敗したりする可能性があります。

DynamoDB から Amazon Simple Storage Service (Amazon S3) または HDFS にバイナリ型をエクスポートすると、Base64 のエンコード文字列として格納されます。Amazon S3 または HDFS のデータを DynamoDB のバイナリ型にインポートする場合、Base64 文字列としてエンコードされます。

## Hive のオプション

次の Hive オプションを設定して、Amazon DynamoDB のデータの転送を管理できます。これらのオプションは、現在の Hive セッションの間のみ、有効です。Hive コマンドプロンプトを終了した後で、クラスターでこのプロンプトを再び開いた場合、この設定はデフォルト値に戻ります。

Hive のオプション	説明
<code>dynamodb.throughput.read.percent</code>	<p>DynamoDB のプロビジョニングされたスループットレートが、表の割り当てられた範囲内に収まるように、読み取りオペレーションのレートを設定します。値は 0.1~1.5 以内にします。</p> <p>0.5 の値がデフォルトの読み取りレートです。つまり、Hive は、表に記載されている読み取りのプロビジョニングされた全体のリソースの半分を消費しようと試みます。値を 0.5 から増やすと、読み取りリクエストレートも増えます。0.5 未満に減らすと、読み取りリクエストレートも減ります。この読み取りレートは概算です。実際の読み取りレートは、DynamoDB に統一ディストリビューションのキーがあるかどうかなどの要因によって変わります。</p> <p>Hive 操作によって、プロビジョニングされたスループットを頻繁に超える場合、またはライブの読み取りトラフィックがスロットリングされる回数が多い場合、この値を 0.5 未満に減らします。十分な容量があり、Hive 操作を高速にする場合は、この値を 0.5 より増やします。また、使用できる未使用の入力/出力操作があると考えられる場合、最大 1.5 に設定してオーバーサブスクライブを実行することもできます。</p>
<code>dynamodb.throughput.write.percent</code>	<p>DynamoDB のプロビジョニングされたスループットレートが、表の割り当てられた範囲内に収まるように、書き込みオペレーションのレートを設定します。値は 0.1~1.5 以内にします。</p> <p>0.5 の値がデフォルトの書き込みレートです。つまり、Hive は、表に記載されている書き込みのプロビジョニングされた全体のリソースの半分を消費しようと試みます。値を 0.5 から増やすと、書き込みリクエストレートも増えます。0.5 未満に減らすと、書き込みリクエストレートも減ります。この書き込みレート</p>

Hive のオプション	説明
	<p>は概算です。実際の書き込みレートは、DynamoDB に統一デистриビューションのキーがあるかどうかなどの要因によって変わります。</p> <p>Hive 操作によって、プロビジョニングされたスループットを頻繁に超える場合、またはライブの書き込みトラフィックがスロットリングされる回数が多い場合、この値を 0.5 未満に減らします。十分な容量があり、Hive 操作を高速にする場合は、この値を 0.5 より増やします。また、使用できる未使用の入力/出力操作があると考えられる場合、または、テーブルに対する最初のデータアップロードであり、ライブトラフィックがまだない場合は、最大 1.5 に設定してオーバーサブスクライブを実行することもできます。</p>
dynamodb.endpoint	DynamoDB サービスのエンドポイントを指定します。使用できる DynamoDB エンドポイントの詳細については、「 <a href="#">リージョンとエンドポイント</a> 」を参照してください。
dynamodb.max.map.tasks	DynamoDB からデータを読み取る時は、マップタスクの最大数を指定します。この値は、1 以上にする必要があります。
dynamodb.retry.duration	Hive コマンドの再試行のタイムアウト期間として使用する分数を指定します。この値は、0 以上の整数にする必要があります。デフォルトのタイムアウト期間は 2 分です。

これらのオプションは、次の例に示すように、SET コマンドを使用して設定されます。

```
SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

## DynamoDB 内データのエクスポート、インポート、クエリを行う Hive コマンドの使用例

以下の例では、Hive コマンドを使用して、データを Amazon S3 や HDFS にエクスポートする、データを DynamoDB にインポートする、テーブルを結合する、テーブルにクエリを行うなどの操作を行います。

Hive テーブルに対する操作では、DynamoDB に格納されているデータを参照します。Hive コマンドは、DynamoDB テーブルに設定されたスループット設定の制約を受けます。また、取得されるデータには、Hive 操作リクエストが DynamoDB で処理された時点で DynamoDB テーブルに書き込まれているデータが含まれます。データ取得プロセスに時間がかかる場合、Hive コマンドによって返されたデータには、Hive コマンドの開始後に DynamoDB で更新されたものが含まれる可能性があります。

Hive コマンドの DROP TABLE と CREATE TABLE は、Hive のローカルテーブルにのみ作用し、DynamoDB のテーブルの作成または削除は行いません。Hive クエリが DynamoDB のテーブルを参照する場合、そのテーブルは、クエリを実行する前に存在している必要があります。DynamoDB でのテーブルの作成と削除の詳細については、「Amazon DynamoDB デベロッパーガイド」の「[DynamoDB のテーブルの操作](#)」を参照してください。

### Note

Hive テーブルを Amazon S3 内の場所にマッピングする場合、バケットのルートパス (s3://mybucket) にはマッピングしないでください。これを行うと、Hive がデータを Amazon S3 に書き込むときにエラーが発生することがあります。代わりに、テーブルをバケットのサブパス ( s3://mybucket/mypath ) にマッピングします。

## DynamoDB からデータをエクスポートする

Hive を使用して DynamoDB のデータをエクスポートすることができます。

DynamoDB テーブルを Amazon S3 バケットにエクスポートするには

- DynamoDB に格納されたデータを参照する Hive テーブルを作成します。次に、INSERT OVERWRITE コマンドを呼び出して、データを外部ディレクトリに書き込みます。次の例で

は、`s3://bucketname/path/subpath/` は Amazon S3. 内の有効なパスです。CREATE コマンドで列とデータ型が DynamoDB 内の値と一致するように調整します。これを使用して DynamoDB データのアーカイブを Amazon S3 に作成できます。

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodhtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE DIRECTORY 's3://bucketname/path/subpath/' SELECT *
FROM hiveTableName;
```

書式設定を使用して DynamoDB テーブルを Amazon S3 バケットにエクスポートするには

- Amazon S3 内の場所を参照する外部テーブルを作成します。これは次の例では `s3_export` です。CREATE の呼び出しで、テーブルの行の書式を指定します。次に、INSERT OVERWRITE を使用して DynamoDB のデータを `s3_export` にエクスポートすると、データは指定の書式で書き込まれます。次の例では、データはカンマ区切り値 ( CSV ) 形式で書き込まれます。

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodhtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

列のマッピングを指定せずに DynamoDB テーブルを Amazon S3 バケットにエクスポートするには

- DynamoDB に格納されたデータを参照する Hive テーブルを作成します。これは、カラムマッピングを指定しないこと以外、前の例とほぼ同じです。このテーブルには `map<string, string>` 型の列が 1 つだけ含まれている必要があります。次に、Amazon S3 に EXTERNAL テーブルを作成したら、INSERT OVERWRITE コマンドを呼び出して DynamoDB のデータを Amazon S3 に書き込みます。これを使用して DynamoDB データのアーカイブを Amazon S3 に作成できます。カラムマッピングがないので、この方法でエクスポートされたテーブルをクエリすることはできません。カラムマッピングを指定しないデータのエクスポートは、Amazon EMR AMI 2.2.x 以降でサポートされる Hive 0.8.1.5 以降で利用できます。

```
CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");

CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE s3TableName SELECT *
FROM hiveTableName;
```

データ圧縮を使用して DynamoDB テーブルを Amazon S3 バケットにエクスポートするには

- Hive では、Hive セッション中に設定できる圧縮コーデックが複数あります。これを行うことで、エクスポートデータは、指定した形式で圧縮されます。次の例では、LZO (Lempel-Ziv-Oberhumer) アルゴリズムを使用して、エクスポートされたファイルを圧縮します。

```
SET hive.exec.compress.output=true;
SET io.seqfile.compression.type=BLOCK;
SET mapred.output.compression.codec = com.hadoop.compression.lzo.LzopCodec;

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
```

```
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

CREATE EXTERNAL TABLE lzo_compression_table (line STRING)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE lzo_compression_table SELECT *
FROM hiveTableName;
```

以下の圧縮コーデックを利用できます。

- org.apache.hadoop.io.compress。GzipCodec
- org.apache.hadoop.io.compress。DefaultCodec
- com.hadoop.compression.lzo。LzoCodec
- com.hadoop.compression.lzo。LzopCodec
- org.apache.hadoop.io.compress.BZip2Codec
- org.apache.hadoop.io.compress。SnappyCodec

DynamoDB テーブルを HDFS にエクスポートするには

- 次の Hive コマンドを使用します。ここで、*hdfs:///directoryName* は有効な HDFS パスであり、*hiveTableName* は DynamoDB を参照する Hive のテーブルです。Hive 0.7.1.1 はデータを Amazon S3 にエクスポートする際に HDFS を中間ステップとして使用するので、このエクスポート操作は、DynamoDB テーブルを Amazon S3 にエクスポートするよりも高速です。また、次の例は、`dynamodb.throughput.read.percent` を 1.0 に設定して読み取りリクエストレートを上げる方法を示しています。

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodhtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE DIRECTORY 'hdfs:///directoryName' SELECT * FROM hiveTableName;
```



Amazon S3 へのエクスポートで説明したように、書式設定や圧縮を使用して HDFS にデータをエクスポートすることもできます。これを行うには、上記の例の Amazon S3 ディレクトリを HDFS ディレクトリに置き換えるだけです。

印刷不可の UTF-8 文字データを Hive で読み取るには

- テーブル作成時に `STORED AS SEQUENCEFILE` 句を使用すると、印刷不可の UTF-8 文字データを Hive で読み取ることができます。SequenceFile は Hadoop バイナリファイル形式です。このファイルを読み取るには Hadoop を使用する必要があります。次の例に、DynamoDB から Amazon S3 にデータをエクスポートする方法を示します。この機能を使用して印刷不可の UTF-8 でエンコードされた文字を処理できます。

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodhtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
STORED AS SEQUENCEFILE
LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

## DynamoDB へのデータのアップロード

Hive を使用してデータを DynamoDB に書き込む場合は、書き込み容量単位の数をクラスター内のマッパの数より大きいことを確認する必要があります。たとえば、m1.xlarge EC2 インスタンスで実行されるクラスターは、インスタンスごとに 8 個のマッパを作成します。10 個のインスタンスがあるクラスターの場合は、合計 80 個のマッパがあることとなります。書き込み容量単位がクラスター内のマッパの数以下である場合、Hive の書き込み操作は書き込みスループットを消費しきるか、プロビジョニングされた以上のスループットを消費しようとしています。各 EC2 インスタンスタイプによって生成されるマッパの数の詳細については、[Hadoop の設定](#) を参照してください。



Hadoop 内のマッパの数は入力分割数によって決まります。分割数が少なすぎる場合、書き込みコマンドは書き込みスループットを消費しきれない可能性があります。

同じキーを持つ項目がターゲット DynamoDB テーブルに存在する場合、項目は上書きされます。キーを持つ項目がターゲット DynamoDB テーブルに存在しない場合、その項目は挿入されます。

Amazon S3 のテーブルを DynamoDB にインポートするには

- Amazon EMR (Amazon EMR) と Hive を使用して、Amazon S3 から DynamoDB にデータを書き込むことができます。

```
CREATE EXTERNAL TABLE s3_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://bucketname/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM s3_import;
```

カラムマッピングを指定せずに Amazon S3 バケットのテーブルを DynamoDB にインポートするには

- 以前 DynamoDB からエクスポートされ Amazon S3 に格納されたデータを参照する EXTERNAL テーブルを作成します。インポートする前に、テーブルが DynamoDB に存在することと、そのキースキーマが、以前エクスポートされた DynamoDB テーブルと同じであることを確認します。また、テーブルには map<string, string> 方の列が 1 つだけ含まれる必要があります。次に、DynamoDB にリンクされている Hive テーブルを作成したら、INSERT OVERWRITE コマンドを呼び出して Amazon S3 のデータを DynamoDB に書き込みます。カラムマッピングがないので、この方法でインポートされたテーブルにクエリを行うことはできません。カラムマッピングを指定しないデータのインポートは、Amazon EMR AMI 2.2.3 以降でサポートされる Hive 0.8.1.5 以降で利用できます。

```
CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
```

```
LOCATION 's3://bucketname/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodhtable1");

INSERT OVERWRITE TABLE hiveTableName SELECT *
FROM s3TableName;
```

HDFS のテーブルを DynamoDB にインポートするには

- Amazon EMR と Hive を使用して HDFS のデータを DynamoDB にインポートすることができません。

```
CREATE EXTERNAL TABLE hdfs_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 'hdfs:///directoryName';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodhtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM hdfs_import;
```

## DynamoDB 内データのクエリ

次の例は、Amazon EMR を使用して DynamoDB のデータにクエリを行うさまざまな方法を示します。

マッピングされたカラムで最大値を検索するには ( max )

- 次のような Hive コマンドを使用します。最初のコマンドの CREATE ステートメントで、DynamoDB に格納されたデータを参照する Hive テーブルを作成します。次に、SELECT ステートメントで、そのテーブルを使用して、DynamoDB に格納されたデータのクエリを行います。次の例では、指定された顧客による最大の注文を検索します。

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,  
items_purchased array<String>)  
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'  
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",  
"dynamodb.column.mapping" =  
"customerId:CustomerId,total_cost:Cost,items_purchased:Items");  
  
SELECT max(total_cost) from hive_purchases where customerId = 717;
```

### GROUP BY 句を使用してデータを集計するには

- GROUP BY 句を使用して、複数のレコードのデータを収集できます。多くの場合、これは sum、count、min、または max のような集計関数とともに使用されます。次の例は、4 件以上注文した顧客の注文のうち最大の注文のリストを返します。

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,  
items_purchased array<String>)  
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'  
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",  
"dynamodb.column.mapping" =  
"customerId:CustomerId,total_cost:Cost,items_purchased:Items");  
  
SELECT customerId, max(total_cost) from hive_purchases GROUP BY customerId HAVING  
count(*) > 3;
```

### 2 つの DynamoDB テーブルを結合するには

- 次の例では、2 つの Hive テーブルを DynamoDB に格納されているデータにマッピングします。その後、2 つのテーブルに対して join を呼び出します。結合はクラスターで計算され、以下を返します。結合は DynamoDB 内では発生しません。この例は、3 件以上注文した顧客について、顧客とその購入品のリストを返します。

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,  
items_purchased array<String>)  
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'  
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",  
"dynamodb.column.mapping" =  
"customerId:CustomerId,total_cost:Cost,items_purchased:Items");  
  
CREATE EXTERNAL TABLE hive_customers(customerId bigint, customerName string,  
customerAddress array<String>)  
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'  
TBLPROPERTIES ("dynamodb.table.name" = "Customers",  
"dynamodb.column.mapping" =  
"customerId:CustomerId,customerName:Name,customerAddress:Address");  
  
Select c.customerId, c.customerName, count(*) as count from hive_customers c  
JOIN hive_purchases p ON c.customerId=p.customerId  
GROUP BY c.customerId, c.customerName HAVING count > 2;
```

異なるソースの2つのテーブルを結合するには

- 次の例の `Customer_S3` は、Amazon S3 に格納された CSV ファイルを読み込む Hive テーブルであり、`hive_purchases` は、DynamoDB のデータを参照するテーブルです。次の例では、Amazon S3 に CSV ファイルとして格納されている顧客データと、DynamoDB に格納されている注文データを結合し、名前に「Miller」が含まれる顧客による注文を表すデータのセットを返します。

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,  
items_purchased array<String>)  
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'  
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",  
"dynamodb.column.mapping" =  
"customerId:CustomerId,total_cost:Cost,items_purchased:Items");  
  
CREATE EXTERNAL TABLE Customer_S3(customerId bigint, customerName string,  
customerAddress array<String>)  
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','  
LOCATION 's3://bucketname/path/subpath/';  
  
Select c.customerId, c.customerName, c.customerAddress from
```

```
Customer_S3 c
JOIN hive_purchases p
ON c.customerid=p.customerid
where c.customerName like '%Miller%';
```

### Note

前の各例には、明確かつ完全にするために CREATE TABLE 句が含まれていました。指定された Hive テーブルに対して複数のクエリやエクスポート操作を実行する場合は、Hive セッションの開始時にテーブルを一度作成するだけで済みます。

## DynamoDB での Amazon EMR オペレーションのパフォーマンスの最適化

DynamoDB テーブルでの Amazon EMR の操作は読み込み操作としてカウントされ、テーブルのプロビジョニング済みスループットの設定が適用されます。Amazon EMR は独自のロジックを実装して、DynamoDB テーブルの負荷を分散させ、プロビジョニング済みスループットを超過する可能性を最小限に抑えようとしています。Amazon EMR はそれぞれの Hive クエリの最後に、プロビジョニング済みスループットを超過した回数など、クエリの処理に使用されたクラスターに関する情報を返します。この情報と DynamoDB スループットに関する CloudWatch メトリクスを使用して、後続のリクエストで DynamoDB テーブルの負荷をより適切に管理できます。

DynamoDB テーブルを操作する際に Hive クエリのパフォーマンスに影響を与える要因を次に示します。

### プロビジョニングされた読み込み容量単位

DynamoDB テーブルに対して Hive クエリを実行する際には、十分な量の読み込み容量単位をプロビジョニングしておく必要があります。

たとえば、DynamoDB テーブルに対して 100 ユニットの読み込みキャパシティーをプロビジョニングしているとします。この場合、1 秒間に 100 の読み込み ( 409,600 バイト ) を実行できます。そのテーブルに 20 GB ( 21,474,836,480 バイト ) のデータが含まれており、Hive クエリがフルテーブルスキャンを実行する場合、クエリの実行にかかる時間は次のように見積もられます。

$$21,474,836,480 / 409,600 = 52,429 \text{ 秒} = 14.56 \text{ 時間}$$

必要な時間を短縮するには、ソース DynamoDB テーブルで読み込みキャパシティーユニットを調整する以外に方法はありません。Amazon EMR クラスターにノードをさらに追加しても、時間は短縮されません。

Hive 出力では、1 つ以上のマッパープロセスが終了すると、完了のパーセンテージが更新されます。プロビジョニングされた読み込みキャパシティーが小さく設定された大きな DynamoDB テーブルでは、完了のパーセンテージ出力が長時間更新されない場合があります。上記のような場合、ジョブは数時間にわたって 0% 完了として表示されます。ジョブの進行状況の詳細なステータスについては、Amazon EMR コンソールに移動してください。ここで、個別のマッパータスクのステータスおよびデータ読み込みの統計を表示できます。

また、マスターノードの Hadoop インターフェイスにログオンし、Hadoop 統計を表示することもできます。ここには、個別のマッパータスクのステータスおよびいくつかのデータ読み込み統計が表示されます。詳細については、「Amazon EMR 管理ガイド」の「[マスターノードでホストされているウェブインターフェイス](#)」を参照してください。

## 読み込みパーセントの設定

デフォルトでは、Amazon EMR は現在のプロビジョニング済みスループットに基づいて、DynamoDB テーブルに対するリクエスト負荷を管理します。ただし、プロビジョニング済みスループットを超過した応答を多数含むジョブに関する情報を Amazon EMR が返す場合は、Hive テーブルを設定する際に、`dynamodb.throughput.read.percent` パラメータを使用してデフォルトの読み込みレートを調整することができます。読み込みパーセントパラメータの設定の詳細については、「[Hive のオプション](#)」を参照してください。

## 書き込みパーセントの設定

デフォルトでは、Amazon EMR は現在のプロビジョニング済みスループットに基づいて、DynamoDB テーブルに対するリクエスト負荷を管理します。ただし、プロビジョニング済みスループットを超過した応答を多数含むジョブに関する情報を Amazon EMR が返す場合は、Hive テーブルを設定する際に、`dynamodb.throughput.write.percent` パラメータを使用してデフォルトの書き込みレートを調整することができます。書き込みパーセントパラメータの設定の詳細については、「[Hive のオプション](#)」を参照してください。

## 再試行間隔の設定

デフォルトでは、Hive クエリが 2 分以内 (デフォルトの再試行間隔) に結果を返さない場合、Amazon EMR はこのクエリを再実行します。この間隔は、Hive クエリの実行時に `dynamodb.retry.duration` パラメータを設定することによって調整できます。書き込みパーセントパラメータの設定の詳細については、「[Hive のオプション](#)」を参照してください。

## マップタスクの数

DynamoDB に格納されているデータのエクスポートおよびクエリ実行のリクエストを処理するために Hadoop が起動するマッパーデーモンは、最大読み込みレート (毎秒 1 MiB) を上限値として、使用される読み込みキャパシティーを制限します。DynamoDB で追加のプロビジョニング済みスループットを使用できる場合は、マッパーデーモンの数を増やせば、Hive エクスポートおよびクエリオペレーションのパフォーマンスを向上させることができます。そのためには、クラスター内の EC2 インスタンスの数を増やすか、または各 EC2 インスタンスで実行されているマッパーデーモンの数を増やします。

クラスター内の EC2 インスタンスの数を増やすには、現在のクラスターを停止し、より多くの EC2 インスタンスとともに再起動します。Amazon EMR コンソールからクラスターを起動する場合は、[Configure EC2 Instances] (EC2 インスタンスの設定) ダイアログボックスで EC2 インスタンスの数を指定します。または、CLI からクラスターを起動する場合は `--num-instances` オプションを使用します。

インスタンスで実行されるマップタスクの数は、EC2 インスタンスタイプによって異なります。サポートされる EC2 インスタンスタイプ、および各タイプで提供されるマッパーの数の詳細については、[タスクの設定](#) を参照してください。そこには、サポートされる設定ごとに "タスクの設定" セクションがあります。

マッパーデーモンの数を増やすもう 1 つの方法としては、Hadoop の `mapreduce.tasktracker.map.tasks.maximum` 設定パラメータをより大きい値に変更します。この場合は、EC2 インスタンスの数またはサイズを大きくせずにマッパーの数を増やすことができるので、コストを削減できるという利点があります。欠点としては、この値を大きく設定しすぎると、クラスター内の EC2 インスタンスがメモリ不足になる可能性があります。`mapreduce.tasktracker.map.tasks.maximum` を設定するには、クラスターを起動し、`mapreduce.tasktracker.map.tasks.maximum` の値を `mapred-site` 設定分類のプロパティとして指定します。以下の例ではこれを示しています。詳細については、「[アプリケーションの設定](#)」を参照してください。

```
{
  "configurations": [
    {
      "classification": "mapred-site",
      "properties": {
        "mapred.tasktracker.map.tasks.maximum": "10"
      }
    }
  ]
}
```



```
}
```

## 並列データリクエスト

複数のユーザーまたは複数のアプリケーションから単一のテーブルに複数のデータリクエストが行われると、プロビジョニング済み読み込みスループットが減少し、パフォーマンス速度が低下します。

## 処理間隔

DynamoDB におけるデータの整合性は、各ノードで実行される読み込み/書き込みオペレーションの順序に影響を受けます。Hive クエリの進行中に、別のアプリケーションが DynamoDB テーブルに新しいデータをロードしたり、既存のデータの変更や削除を行ったりする場合があります。この場合、クエリの実行中にデータに対して行われた変更は Hive クエリの結果に反映されないことがあります。

## スループットの超過の回避

DynamoDB に対して Hive クエリを実行する際には、プロビジョニング済みスループットを超過しないように注意してください。超過すると、アプリケーションによる DynamoDB::Get の呼び出しに必要な容量が不足してしまいます。これが発生しないようにするには、Amazon のログとモニタリングメトリクスをチェック DynamoDB::Get して、他のアプリケーション呼び出しの読み取りポリシーとスロットリングを定期的にモニタリングする必要があります CloudWatch。

## [Request time] (リクエストタイム)

DynamoDB テーブルの需要が低いタイミングでテーブルにアクセスするように、Hive クエリをスケジューリングすると、パフォーマンスを向上させられます。たとえば、アプリケーションのほとんどのユーザーがサンフランシスコに住んでいる場合、大部分のユーザーが睡眠中で DynamoDB データベースを更新していない毎朝 4 時 (PST) にデータをエクスポートするように選択することができます。

## 時間ベースのテーブル

データが一連の時間ベースの DynamoDB テーブル (たとえば、1 日あたり 1 つのテーブル) として構成されている場合は、テーブルが非アクティブになったときにデータをエクスポートできます。この手法を使用すると、データを Amazon S3 に継続的にバックアップできます。

## アーカイブされたデータ

DynamoDB に格納されているデータに対して多数の Hive クエリを実行する予定であり、アーカイブされたデータがアプリケーションで許容される場合は、データを HDFS または Amazon S3 にエクス



ポートし、DynamoDB の代わりにデータのコピーに対して Hive クエリを実行することができます。これにより、読み込みオペレーションおよびプロビジョニング済みスループットを過度に使用せずに済みます。

## Kinesis

Amazon EMR クラスターは、Hive、Pig、Hadoop Streaming API、カスケードなどの Hadoop エコシステムの使い慣れたツールを使用して、Amazon Kinesis MapReduce ストリームを直接読み取って処理できます。また、実行しているクラスターで Amazon Kinesis のリアルタイムデータを Amazon S3、Amazon DynamoDB、および HDFS の既存データと結合できます。後処理のアクティビティとして、Amazon EMR から Amazon S3 または DynamoDB に直接データをロードできます。Amazon Kinesis の主なサービスと料金については、「[Amazon Kinesis](#)」ページを参照してください。

## Amazon EMR と Amazon Kinesis の統合で何ができますか？

Amazon EMR と Amazon Kinesis の統合により、以下のようなシナリオへの対応が非常に容易になります。

- ストリーミングログ分析 – ストリーミングのウェブログを分析して、リージョン別、ブラウザ別、およびアクセスドメイン別に、数分ごとの上位 10 件のエラータイプのリストを生成できます。
- カスタマーエンゲージメント – Amazon Kinesis のクリックストリームデータと DynamoDB テーブルに保存されている広告キャンペーン情報を結合するクエリを作成し、特定のウェブサイトに表示される最も効果的な広告カテゴリを特定できます。
- アドホックインタラクティブクエリ – Amazon Kinesis Streams から HDFS に定期的にデータを読み込み、ローカルの Impala テーブルとして使用可能にすることで、高速かつインタラクティブな分析クエリを実行できます。

## Amazon Kinesis Streams のチェックポイントの分析

ユーザーは、Amazon Kinesis Streams の定期的なバッチ分析をいわゆる反復で実行できます。Amazon Kinesis Streams のデータレコードはシーケンス番号を使用して取得されるため、反復の境界は Amazon EMR によって DynamoDB テーブルに格納される開始と終了のシーケンス番号で定義されます。たとえば、iteration0 が終了すると、終了のシーケンス番号が DynamoDB に格納されるため、iteration1 ジョブが開始されたとき、ストリームからそれに続くデータを取得できます。このストリームデータの反復のマッピングはチェックポイントと呼ばれます。詳細については、「[Kinesis コネクタ](#)」を参照してください。

反復にチェックポイントが設定された後に、ジョブの反復処理が失敗した場合、Amazon EMR では、その反復のレコード処理が再試行されます。

チェックポイントは、次のことが可能になる機能です。

- 同じストリームと論理名で実行した前のクエリにより処理された連続番号の後にデータ処理を開始します
- 前のクエリで処理された Kinesis のデータと同じバッチを再処理します

チェックポイントを有効にするには、スクリプトで `kinesis.checkpoint.enabled` パラメータを `true` に設定します。また、以下のパラメータを設定します。

構成設定	説明
<code>kinesis.checkpoint.metastore.table.name</code>	チェックポイント情報が保存される DynamoDB テーブル名
<code>kinesis.checkpoint.metastore.hash.key.name</code>	DynamoDB テーブルのハッシュキー名
<code>kinesis.checkpoint.metastore.hash.range.name</code>	DynamoDB テーブルの範囲キー名
<code>kinesis.checkpoint.logical.name</code>	現在の処理の論理名
<code>kinesis.checkpoint.iteration.no</code>	論理名に関連付けられている処理の反復番号
<code>kinesis.rerun.iteration.without.wait</code>	失敗した反復がタイムアウトを待たずに再実行できるかどうかを示すブール値 ( デフォルトは <code>false</code> )。

## Amazon DynamoDB テーブルのプロビジョニングされた IOPS に関する推奨事項

Amazon Kinesis 用の Amazon EMR コネクタは、チェックポイントのメタデータの補助として DynamoDB データベースを使用します。Amazon Kinesis Streams のデータをチェックポイントの間隔で Amazon EMR クラスターで使用する前に、DynamoDB のテーブルを作成する必要があります。テーブルは Amazon EMR クラスターと同じリージョンに存在する必要があります。以下は、DynamoDB テーブルにプロビジョニングする必要がある IOPS の数に関する一般的な推奨で

す。j は同時に実行できる Hadoop ジョブ (異なる論理名 + 反復数の組み合わせ) の最大数で、s はジョブが処理するシャードの最大数です。

読み込みキャパシティーユニット:  $j*s/5$

書き込みキャパシティーユニット:  $j*s$

## パフォーマンスに関する考慮事項

Amazon Kinesis シャードスループットは、Amazon EMR クラスターのノードのインスタンスのサイズ、およびストリームのレコードのサイズに正比例しています。マスターノードやコアノードで m5.xlarge がそれ以上のインスタンスを使用することをお勧めします。

## Amazon EMR で Amazon Kinesis 分析をスケジュールする

任意の繰り返しについてタイムアウトと最大期間で制限される、アクティブな Amazon Kinesis Streams でデータを分析するときは、ストリームから定期的に詳細を取得するために、分析を頻繁に実行することが重要です。定期的な間隔でこのようなスクリプトおよびクエリを実行する方法は複数ありますが、これらのような反復タスクには、AWS Data Pipeline を使用することをお勧めします。詳細については、「AWS Data Pipeline デベロッパーガイド [AWS Data Pipeline HiveActivity](#)」の [AWS Data Pipeline PigActivity](#) 「」および「」を参照してください。

## Amazon EMR 7.0 用スパークキネシスコネクターの SDK 2.x への移行

AWS SDK は、認証情報の管理、S3 および Kinesis サービスへの接続など、クラウドコンピューティングサービスとやり取り AWS するための豊富な APIs とライブラリのセットを提供します。Spark Kinesis コネクタは Kinesis データストリームからデータを消費するために使用され、受信したデータは Spark の実行エンジンで変換および処理されます。現在、このコネクタは AWS SDK と Kinesis-client-library (KCL) の 1.x 上に構築されています。

AWS SDK 2.x の移行の一環として、Spark Kinesis コネクタも SDK 2.x で実行されるように更新されます。Amazon EMR 7.0 リリースの Spark には、コミュニティバージョンの Apache Spark ではまだ利用できない SDK 2.x アップグレードが含まれています。7.0 より前のリリースの Spark Kinesis コネクタを使用する場合、Amazon EMR 7.0 に移行する前に、アプリケーションコードを SDK 2.x で実行するように移行する必要があります。

## 移行ガイド

このセクションでは、アップグレードされた Spark Kinesis コネクタにアプリケーションを移行する手順について説明します。これには、AWS SDK 2.x の Kinesis Client Library (KCL) 2.x、AWS 認証

情報プロバイダー、および AWS サービスクライアントに移行するためのガイドが含まれています。参考までに、Kinesis コネクタを使用するサンプル[WordCount](#)プログラムも含まれています。

## トピック

- [1.x から 2.x への移行](#)
- [AWS SDK 1.x から 2.x への AWS 認証情報プロバイダーの移行](#)
- [AWS SDK 1.x から 2.x への AWS サービスクライアントの移行](#)
- [ストリーミングアプリケーションのコード例](#)
- [アップグレードされた Spark Kinesis コネクタを使用する際の考慮事項](#)

## 1.x から 2.x への移行

- **KinesisInputDStream** のメトリクスレベルとディメンション

KinesisInputDStream をインスタンス化すると、ストリームのメトリクスレベルとディメンションを制御できます。以下の例は、KCL 1.x でこれらのパラメータをカスタマイズする方法を示しています。

```
import
  com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel

val kinesisStream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(kinesisCheckpointInterval)
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
  .metricsLevel(MetricsLevel.DETAILED)

  .metricsEnabledDimensions(KinesisClientLibConfiguration.DEFAULT_METRICS_ENABLED_DIMENSIONS.a
  .build()
```

KCL 2.x では、これらの構成設定のパッケージ名は異なります。2.x に移行するには:

1. `com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibCon` および `com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel` のインポートステートメントをそれぞれ `software.amazon.kinesis.metrics.MetricsLevel` および `software.amazon.kinesis.metrics.MetricsUtil` に変更します。

```
// import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel
import software.amazon.kinesis.metrics.MetricsLevel

// import
com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
import software.amazon.kinesis.metrics.MetricsUtil
```

2. 行 `metricsEnabledDimensionsKinesisClientLibConfiguration.DEFAULT_METRICS_ENABLED` を `metricsEnabledDimensionsSet(MetricsUtil.OPERATION_DIMENSION_NAME, MetricsUtil.SHARD_ID_DIMENSION_NAME)` に置き換えます。

以下は、カスタマイズされたメトリクスレベルとメトリクスディメンションを備えた `KinesisInputDStream` の更新バージョンです。

```
import software.amazon.kinesis.metrics.MetricsLevel
import software.amazon.kinesis.metrics.MetricsUtil

val kinesisStream = KinesisInputDStream.builder
    .streamingContext(ssc)
    .streamName(streamName)
    .endpointUrl(endpointUrl)
    .regionName(regionName)
    .initialPosition(new Latest())
    .checkpointAppName(appName)
    .checkpointInterval(kinesisCheckpointInterval)
    .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
    .metricsLevel(MetricsLevel.DETAILED)
    .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
    MetricsUtil.SHARD_ID_DIMENSION_NAME))
    .build()
```

- `KinesisInputDStream` のメッセージハンドラー関数

KinesisInputDStream をインスタンス化するときに、パーティションキーなどの Record に含まれる他のデータを使用したい場合に備えて、Kinesis Record を取得して汎用オブジェクト T を返す「メッセージハンドラー関数」を提供することもできます。

KCL 1.x では、メッセージハンドラー関数のシグネチャは Record => T であり、レコードは `com.amazonaws.services.kinesis.model.Record` です。KCL 2.x では、ハンドラーの署名はに変更されます。ここで `KinesisClientRecord => T KinesisClientRecord`、は `software.amazon.kinesis.retrieval.KinesisClientRecord`。

KCL 1.x でメッセージハンドラーを提供する例を以下に示します。

```
import com.amazonaws.services.kinesis.model.Record

def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

メッセージハンドラーを移行するには:

1. `com.amazonaws.services.kinesis.model.Record` のインポートステートメントを `software.amazon.kinesis.retrieval.KinesisClientRecord` に変更します。

```
// import com.amazonaws.services.kinesis.model.Record
import software.amazon.kinesis.retrieval.KinesisClientRecord
```

2. メッセージハンドラーのメソッドシグネチャを更新します。

```
//def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt + 5
```

KCL 2.x でメッセージハンドラーを提供する、更新済みの例を以下に示します。

```
import software.amazon.kinesis.retrieval.KinesisClientRecord

def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

KCL 1.x から 2.x への移行について詳しくは、「[KCL 1.x から KCL 2.x へのコンシューマの移行](#)」を参照してください。

## AWS SDK 1.x から 2.x への AWS 認証情報プロバイダーの移行

認証情報プロバイダーは、とのやり取りの AWS 認証情報を取得するために使用されます AWS。SDK 2.x の認証情報プロバイダーに関連するインターフェイスとクラスの変更がいくつかあります。詳しくは、[こちら](#)をご覧ください。Spark Kinesis コネクタは、AWS 認証情報プロバイダーの 1.x バージョンを返すインターフェイス (`org.apache.spark.streaming.kinesis.SparkAWSCredentials`) と実装クラスを定義しています。これらの認証情報プロバイダーは、Kinesis クライアントを初期化する際に必要です。例えば、`SparkAWSCredentials.provider` アプリケーションでメソッドを使用している場合は、AWS 認証情報プロバイダーの 2.x バージョンを使用するようにコードを更新する必要があります。

AWS SDK 1.x で認証情報プロバイダーを使用する例を次に示します。

```
import org.apache.spark.streaming.kinesis.SparkAWSCredentials
import com.amazonaws.auth.AWSCredentialsProvider

val basicSparkCredentials = SparkAWSCredentials.builder
  .basicCredentials("accessKey", "secretKey")
  .build()
```

```
val credentialProvider = basicSparkCredentials.provider
assert(credentialProvider.isInstanceOf[AWSCredentialsProvider], "Type should be
  AWSCredentialsProvider")
```

SDK 2.x に移行するには:

1. `com.amazonaws.auth.AWSCredentialsProvider` のインポートステートメントを `software.amazon.awssdk.auth.credentials.AwsCredentialsProvider` に変更します。

```
//import com.amazonaws.auth.AWSCredentialsProvider
import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider
```

2. このクラスを使用する残りのコードを更新します。

```
import org.apache.spark.streaming.kinesis.SparkAWSCredentials
import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider

val basicSparkCredentials = SparkAWSCredentials.builder
  .basicCredentials("accessKey", "secretKey")
  .build()

val credentialProvider = basicSparkCredentials.provider
assert (credentialProvider.isInstanceOf[AwsCredentialsProvider], "Type should be
  AwsCredentialsProvider")
```

## AWS SDK 1.x から 2.x への AWS サービスクライアントの移行

AWS サービスクライアントは 2.x で異なるパッケージ名を持っていますが (例:

`software.amazon.awssdk`)、SDK 1.x は `com.amazonaws` を使用します。クライアント変更に関する詳細については、[こちら](#)を参照してください。コードでこれらのサービスクライアントを使用している場合は、それに応じてクライアントを移行する必要があります。

SDK 1.x でクライアントを作成する例を以下に示します。

```
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
import com.amazonaws.services.dynamodbv2.document.DynamoDB

AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
```



```
AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();
```

2.x に移行するには:

1. サービスクライアントのインポートステートメントを変更します。DynamoDB クライアントを例にとってみましょう。  
com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient または com.amazonaws.services.dynamodbv2.document.DynamoDB を software.amazon.awssdk.services.dynamodb.DynamoDbClient に変更する必要があります。

```
// import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
// import com.amazonaws.services.dynamodbv2.document.DynamoDB
import software.amazon.awssdk.services.dynamodb.DynamoDbClient
```

2. クライアントを初期化するコードを更新してください

```
// AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
// AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();

DynamoDbClient ddbClient = DynamoDbClient.create();
DynamoDbClient ddbClient = DynamoDbClient.builder().build();
```

AWS SDK を 1.x から 2.x に移行する方法の詳細については、[AWS 「SDK for Java 1.x と 2.x の違い」を参照してください。](#)

ストリーミングアプリケーションのコード例

```
import java.net.URI
import software.amazon.awssdk.auth.credentials.DefaultCredentialsProvider
import software.amazon.awssdk.http.apache.ApacheHttpClient
import software.amazon.awssdk.services.kinesis.KinesisClient
import software.amazon.awssdk.services.kinesis.model.DescribeStreamRequest
import software.amazon.awssdk.regions.Region
import software.amazon.kinesis.metrics.{MetricsLevel, MetricsUtil}

import org.apache.spark.SparkConf
import org.apache.spark.storage.StorageLevel
import org.apache.spark.streaming.{Milliseconds, StreamingContext}
import org.apache.spark.streaming.dstream.DStream.toPairDStreamFunctions
```

```
import org.apache.spark.streaming.kinesis.KinesisInitialPositions.Latest
import org.apache.spark.streaming.kinesis.KinesisInputDStream

object KinesisWordCountASLSDKV2 {

  def main(args: Array[String]): Unit = {
    val appName = "demo-app"
    val streamName = "demo-kinesis-test"
    val endpointUrl = "https://kinesis.us-west-2.amazonaws.com"
    val regionName = "us-west-2"

    // Determine the number of shards from the stream using the low-level Kinesis
    Client
    // from the AWS Java SDK.
    val credentialsProvider = DefaultCredentialsProvider.create
    require(credentialsProvider.resolveCredentials() != null,
      "No AWS credentials found. Please specify credentials using one of the methods
    specified " +
      "in https://docs.aws.amazon.com/sdk-for-java/latest/developer-guide/
    credentials.html")
    val kinesisClient = KinesisClient.builder()
      .credentialsProvider(credentialsProvider)
      .region(Region.US_WEST_2)
      .endpointOverride(URI.create(endpointUrl))
      .httpClientBuilder(ApacheHttpClient.builder())
      .build()
    val describeStreamRequest = DescribeStreamRequest.builder()
      .streamName(streamName)
      .build()
    val numShards = kinesisClient.describeStream(describeStreamRequest)
      .streamDescription
      .shards
      .size

    // In this example, we are going to create 1 Kinesis Receiver/input DStream for
    each shard.
    // This is not a necessity; if there are less receivers/DStreams than the number of
    shards,
    // then the shards will be automatically distributed among the receivers and each
    receiver
    // will receive data from multiple shards.
    val numStreams = numShards
```

```
// Spark Streaming batch interval
val batchInterval = Milliseconds(2000)

// Kinesis checkpoint interval is the interval at which the DynamoDB is updated
with information
// on sequence number of records that have been received. Same as batchInterval for
this
// example.
val kinesisCheckpointInterval = batchInterval

// Setup the SparkConfig and StreamingContext
val sparkConfig = new SparkConf().setAppName("KinesisWordCountASLSDKV2")
val ssc = new StreamingContext(sparkConfig, batchInterval)

// Create the Kinesis DStreams
val kinesisStreams = (0 until numStreams).map { i =>
  KinesisInputDStream.builder
    .streamingContext(ssc)
    .streamName(streamName)
    .endpointUrl(endpointUrl)
    .regionName(regionName)
    .initialPosition(new Latest())
    .checkpointAppName(appName)
    .checkpointInterval(kinesisCheckpointInterval)
    .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
    .metricsLevel(MetricsLevel.DETAILED)
    .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
MetricsUtil.SHARD_ID_DIMENSION_NAME))
    .build()
}

// Union all the streams
val unionStreams = ssc.union(kinesisStreams)

// Convert each line of Array[Byte] to String, and split into words
val words = unionStreams.flatMap(byteArray => new String(byteArray).split(" "))

// Map each word to a (word, 1) tuple so we can reduce by key to count the words
val wordCounts = words.map(word => (word, 1)).reduceByKey(_ + _)

// Print the first 10 wordCounts
wordCounts.print()
```

```
// Start the streaming context and await termination
ssc.start()
ssc.awaitTermination()
}
}
```

## アップグレードされた Spark Kinesis コネクタを使用する際の考慮事項

- アプリケーションが 11 より前のバージョンの JDK の Kinesis-producer-library を使用している場合、`java.lang.NoClassDefFoundError: javax/xml/bind/DatatypeConverter` のような例外が発生する可能性があります。これは、EMR 7.0 にはデフォルトで JDK 17 が付属しており、J2EE モジュールは Java 11 以降で標準ライブラリから削除されているためです。これは pom ファイルに次の依存関係を追加することで修正できます。必要に応じてライブラリバージョンを適切なもので交換します。

```
<dependency>
  <groupId>javax.xml.bind</groupId>
  <artifactId>jaxb-api</artifactId>
  <version>${jaxb-api.version}</version>
</dependency>
```

- EMR クラスターの作成後、Spark Kinesis コネクタ jar はパス `/usr/lib/spark/connector/lib/` にあります。

## S3DistCp (s3-dist-cp)

Apache DistCp は、大量のデータをコピーするために使用できるオープンソースツールです。S3DistCp は に似ていますが DistCp、AWS 特に Amazon S3 と連携するように最適化されています。S3DistCp in Amazon EMR バージョン 4.0 以降のコマンドは `s3-dist-cp`、クラスターまたはコマンドラインにステップとして追加します。S3DistCp を使用すると、大量のデータを Amazon S3 から HDFS に効率的にコピーし、Amazon EMR クラスターの後続のステップで処理できます。S3DistCp を使用して、Amazon S3 バケット間、または HDFS から Amazon S3 にデータをコピーすることもできます。S3DistCp は、バケット間および AWS アカウント間で多数のオブジェクトを並列コピーするための、よりスケーラブルで効率的です。

実際のシナリオにおける S3DistCP の柔軟性を示す特定のコマンドについては、AWS ビッグデータブログの [S3DistCp を使用するための 7 つのヒント](#)」を参照してください。

と同様に DistCp、S3DistCp は MapReduce を使用して分散的にコピーします。複数のサーバー全体でのコピー、エラー処理、復旧、および報告のタスクが共有されます。Apache DistCp オープンソースプロジェクトの詳細については、Apache Hadoop ドキュメント [DistCp](#) のガイドを参照してください。

S3DistCp が指定されたファイルの一部または全部をコピーできない場合、クラスターステップは失敗し、ゼロ以外のエラーコードを返します。この場合、S3DistCp は部分的にコピーされたファイルをクリーンアップしません。

#### Important

S3DistCp は、アンダースコア文字を含む Amazon S3 バケット名をサポートしていません。S3DistCp は Parquet ファイルの連結をサポートしていません。PySpark 代わりにを使用します。詳細については、「[Amazon EMR での Parquet ファイルの連結](#)」を参照してください。

S3DistCp を使用して (ディレクトリではなく) 1 つのファイルを S3 から HDFS にコピーするときにコピーエラーを回避するには、Amazon EMR バージョン 5.33.0 以降、または Amazon EMR 6.3.0 以降を使用してください。

## S3DistCp オプション

に似ていますが DistCp、S3DistCp はデータのコピー方法と圧縮方法を変更するためのさまざまなオプションをサポートしています。

S3DistCp を呼び出すときに、次の表で説明されているオプションを指定できます。オプションは、引数リストを使用してステップに追加されます。S3DistCp 引数の例を次の表に示します。

オプション	説明	必須
<code>--src=LOCATION</code>	コピーするデータのロケーション。HDFS または Amazon S3 ロケーションのいずれかを指定できます。  例: <code>--src=s3://DOC-EXAMPLE-BUCKET1/logs/j-3GYXXXXXX9I0J/node</code>	はい

オプション	説明	必須
	<p><b>⚠ Important</b></p> <p>S3DistCp は、アンダースコア文字を含む Amazon S3 バケット名をサポートしていません。</p>	
<p><code>--dest=LOCATION</code></p>	<p>データのコピー先。HDFS または Amazon S3 ロケーションのいずれかを指定できます。</p> <p>例: <code>--dest=hdfs:///output</code></p> <p><b>⚠ Important</b></p> <p>S3DistCp は、アンダースコア文字を含む Amazon S3 バケット名をサポートしていません。</p>	はい
<p><code>--srcPattern=PATTERN</code></p>	<p><code>--src</code> でのデータのサブセットへのコピー操作をフィルタリングする <a href="#">正規表現</a>。 <code>--srcPattern</code> と <code>--groupBy</code> がどちらも指定されていない場合は、<code>--src</code> のすべてのデータが <code>--dest</code> にコピーされます。</p> <p>正規表現の引数にアスタリスク (*) などの特殊文字が含まれる場合は、正規表現または <code>--args</code> 文字列全体のいずれかを引用符 (') で囲む必要があります。</p> <p>例: <code>--srcPattern=.*daemons.*-hadoop-.*</code></p>	いいえ

オプション	説明	必須
--groupBy=PATTERN	<p>S3DistCp が <a href="#">式</a> に一致するファイルを連結する正規表現。たとえば、このオプションを使用すると、1 時間で書き込まれたすべてのログファイルを 1 つのファイルに統合できます。統合されたファイル名は、グループ化のためのこの正規表現と一致する値になります。</p> <p>括弧は、ファイルをグループ化する方法を示しています。括弧で囲まれたステートメントと一致するすべての項目が 1 つの出力ファイルに統合されます。正規表現に括弧ステートメントが含まれていない場合、クラスターは S3DistCp ステップで失敗し、エラーを返します。</p> <p>正規表現の引数にアスタリスク (*) などの特殊文字が含まれる場合は、正規表現または --args 文字列全体のいずれかを引用符 (') で囲む必要があります。</p> <p>--groupBy が指定されている場合は、指定されたパターンと一致するファイルのみがコピーされます。--groupBy と --srcPattern は同時に指定する必要はありません。</p> <p>例: --groupBy=.*subnetid.*([0-9]+-[0-9]+-[0-9]+).*</p>	いいえ

オプション	説明	必須
<code>--targetSize=SIZE</code>	<p><code>--groupBy</code> オプションに基づいて作成するファイルのサイズ ( メビバイト ( Mib ) 単位 )。整数の値である必要があります。<code>--targetSize</code> が設定されている場合、S3DistCp はこのサイズと一致するように試みます。コピーされたファイルの実際のサイズは、この値よりも大きくなるか小さくなる可能性があります。ジョブはデータファイルのサイズに基づいて集計されるため、ターゲットファイルサイズがソースデータファイルのサイズに一致する場合があります。</p> <p><code>--groupBy</code> によって統合されたファイルが <code>--targetSize</code> の値よりも大きい場合、そのファイルはパートファイルに分割され、それぞれのファイル名の末尾に数値が順番に付けられます。たとえば、<code>myfile.gz</code> に統合されたファイルは、<code>myfile0.gz</code>、<code>myfile1.gz</code> などにパートに分割されます。</p> <p>例: <code>--targetSize=2</code></p>	いいえ
<code>--appendToLastFile</code>	<p>既に存在する Amazon S3DistCp から HDFS にファイルをコピーする際の S3 の動作を指定します。Amazon S3 既存のファイルに新しいファイルのデータが追加されます。<code>--appendToLastFile</code> を指定して <code>--groupBy</code> を使用すると、新しいデータは同じグループのファイルに追加されます。このオプションは、<code>--targetSize</code> を <code>--groupBy</code> を指定して使用した動作にも順守しています。</p>	いいえ



オプション	説明	必須
<code>--outputCodec=CODEC</code>	<p>コピーされたファイルで使用する圧縮コーデックを指定します。値としては <code>gzip</code>、<code>gz</code>、<code>lzo</code>、<code>snappy</code>、または <code>none</code> をとることができます。このオプションを使用すると、たとえば、<code>Gzip</code> で圧縮された入力ファイルを <code>LZO</code> 圧縮の出力ファイルに変換したり、コピーオペレーションの一環としてファイルを解凍したりできます。出力コーデックを選択した場合、ファイル名には適切な拡張子 (<code>gz</code> や <code>gzip</code> など。拡張子は <code>.gz</code>) が付加されます。<code>--outputCodec</code> の値を指定しない場合、圧縮に変更が加えられずにファイルがコピーされます。</p> <p>例: <code>--outputCodec=lzo</code></p>	いいえ
<code>--s3ServerSideEncryption</code>	<p>ターゲットデータが <code>SSL</code> を使用して転送され、<code>AWS</code> サービス側のキーを使用して <code>Amazon S3</code> で自動的に暗号化されるようにします。<code>S3DistCp</code> を使用してデータを取得すると、オブジェクトは自動的に暗号化されません。暗号化されていないオブジェクトを、暗号化が必要な <code>Amazon S3</code> バケットにコピーしようとする、操作は失敗します。詳細については、「<a href="#">データ暗号化の使用</a>」を参照してください。</p> <p>例: <code>--s3ServerSideEncryption</code></p>	いいえ

オプション	説明	必須
<code>--deleteOnSuccess</code>	<p>コピーオペレーションが成功すると、このオプションにより、S3DistCp はコピー元の場所からコピーされたファイルを削除します。このオプションは、スケジュールされたタスクとして、ログファイルなどの出力ファイルを、あるロケーションから他のロケーションにコピーするときに役立ちます。これにより、同じファイルを 2 回コピーせずに済むようになります。</p> <p>例: <code>--deleteOnSuccess</code></p>	いいえ
<code>--disableMultipartUpload</code>	<p>マルチパートアップロードの使用を無効にします。</p> <p>例: <code>--disableMultipartUpload</code></p>	いいえ
<code>--multipartUploadChunkSize=SIZE</code>	<p>Amazon S3 マルチパートアップロードでの各パートサイズ (MiB 単位)。S3DistCp は、より大きいデータをコピーするときにマルチパートアップロードを使用しません <code>multipartUploadChunkSize</code> 。ジョブのパフォーマンスを向上させるには、各パートサイズを大きく設定します。デフォルトのサイズは 128 MiB です。</p> <p>例: <code>--multipartUploadChunkSize=1000</code></p>	いいえ
<code>--numberFiles</code>	<p>出力ファイル名の先頭に通し番号を付けます。 <code>--startingIndex</code> で別の値を指定しない限り、番号は 0 で開始されます。</p> <p>例: <code>--numberFiles</code></p>	いいえ
<code>--startingIndex=INDEX</code>	<p><code>--numberFiles</code> とともに使用され、通し番号の開始値を指定します。</p> <p>例: <code>--startingIndex=1</code></p>	いいえ

オプション	説明	必須
--outputManifest=F ILENAME	<p>S3DistCp によってコピーされたすべてのファイルのリストを含む、Gzip で圧縮されたテキストファイルを作成します。</p> <p>例: --outputManifest=manifest-1.gz</p>	いいえ
--previousManifest=PATH	<p>--outputManifest フラグを使用して S3DistCp への前回の呼び出し中に作成されたマニフェストファイルを読み取ります。--previousManifest フラグが設定されると、S3DistCp はマニフェストにリストされているファイルをコピーオペレーションから除外します。--outputManifest と共に --previousManifest が指定されている場合、新しいマニフェストファイルには、以前のマニフェストにリストされているファイルも表示されます。ただし、そのファイルはコピーされません。</p> <p>例: --previousManifest=/usr/bin/manifest-1.gz</p>	いいえ
--requirePreviousManifest	<p>S3DistCp への前回の呼び出し中に作成された以前のマニフェストが必要です。これが false に設定されている場合、以前のマニフェストが指定されていないときにエラーが生成されません。デフォルトは true です。</p>	いいえ
--copyFromManifest	<p>の動作を逆にして--previousManifest、S3DistCp がコピーから除外するファイルのリストではなく、指定されたマニフェストファイルをコピーするファイルのリストとして使用します。</p> <p>例: --copyFromManifest --previousManifest=/usr/bin/manifest-1.gz</p>	いいえ

オプション	説明	必須
<code>--s3Endpoint=ENDPOINT</code>	<p>ファイルをアップロードするときに使用する Amazon S3 エンドポイントを指定します。このオプションは、アップロード元とアップロード先の両方のエンドポイントを設定します。設定しない場合は、デフォルトのエンドポイント <code>s3.amazonaws.com</code> が使用されます。Amazon S3 エンドポイントの一覧は、「<a href="#">リージョンとエンドポイント</a>」を参照してください。</p> <p>例: <code>--s3Endpoint=s3.eu-west-1.amazonaws.com</code></p>	いいえ
<code>--storageClass=CLASS</code>	<p>送信先が Amazon S3 のときに使用するストレージクラス。有効な値は、STANDARD と REDUCED_REDUNDANCY です。このオプションを指定しない場合、S3DistCp はストレージクラスを保持しようとはしません。</p> <p>例: <code>--storageClass=STANDARD</code></p>	いいえ

オプション	説明	必須
--srcPrefixesFile=PATH	<p>src プレフィックスのリスト (1 行ごとにプレフィックスが 1 つずつ) を含む Amazon S3 (s3://)、HDFS (hdfs://)、またはローカルファイルシステム (file:/) 内のテキストファイル。</p> <p>が指定されている場合、S3DistCp srcPrefixesFile は src パスを一覧表示しません。代わりに、このファイルで指定されたすべてのプレフィックスをリストした結果を組み合わせたソースリストを生成します。これらのプレフィックスの代わりに、src パスに対する相対パスを使用して宛先パスが生成されます。srcPattern も指定された場合、入力をさらにフィルタリングするため、ソースプレフィックスを組み合わせたリスト結果に適用されます。copyFromManifest が使用された場合、マニフェストのオブジェクトがコピーされ、srcPrefixesFile が無視されます。</p> <p>例: --srcPrefixesFile=PATH</p>	いいえ

上記のオプションに加えて、S3DistCp は [ツールインターフェイス](#) を実装します。つまり、汎用オプションがサポートされています。

## S3DistCp をクラスター内のステップとして追加する

クラスターのステップとして追加することで、S3DistCp を呼び出すことができます。ステップは、起動時にクラスターに追加することも、コンソール、CLI、または API を使用して実行中のクラスターに追加することもできます。次の例は、実行中のクラスターに S3DistCp ステップを追加する方法を示しています。クラスターへのステップの追加の詳細については、「Amazon EMR 管理ガイド」の「[クラスターへの作業の送信](#)」を参照してください。

を使用して実行中のクラスターに S3DistCp step を追加するには AWS CLI

での Amazon EMR コマンドの使用の詳細については AWS CLI、[AWS CLI コマンドリファレンス](#) を参照してください。

- S3DistCp を呼び出すクラスターにステップを追加するには、S3DistCp がコピーオペレーションを引数として実行する方法を指定するパラメータを渡します。

次の例では、デーモンログを Amazon S3 から `hdfs:///output` にコピーします。次のコマンドについて説明します。

- `--cluster-id` でクラスターを指定します。
- `Jar` は S3DistCp JAR ファイルの場所です。 `command-runner.jar` を使用してクラスターでコマンドを実行する方法の例については、「[カスタム JAR ステップを送信してスクリプトまたはコマンドを実行する](#)」を参照してください。
- `Args` は、S3DistCp に渡すオプションの名前と値のペアのカンマ区切りリストです。使用可能なオプションの完全なリストについては、「[S3DistCp オプション](#)」を参照してください。

実行中のクラスターに S3DistCp copy ステップを追加するには、`myStep.json` この例のように、Amazon S3 またはローカルファイルシステムに保存された JSON ファイルに以下を入力します。`j-3GYXXXXXX9I0K` をクラスター ID に置き換え、`mybucket` を Amazon S3 バケット名に置き換えます。

```
[
  {
    "Name": "S3DistCp step",
    "Args": ["s3-dist-cp", "--s3Endpoint=s3.amazonaws.com", "--src=s3://mybucket/logs/j-3GYXXXXXX9I0K/node/", "--dest=hdfs:///output", "--srcPattern=.*[a-zA-Z,]+"],
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR",
    "Jar": "command-runner.jar"
  }
]
```

```
aws emr add-steps --cluster-id j-3GYXXXXXX9I0K --steps file:///./myStep.json
```

## Example Amazon S3 からのログファイルを HDFS へコピーする

この例では、実行中のクラスターにステップを追加して、Amazon S3 バケットに格納されたログファイルを HDFS にコピーする方法も示します。この例では、`--srcPattern` オプションを使用して、デーモンログにコピーされたデータを制限しています。

--srcPattern オプションを使用してログファイルを Amazon S3 から HDFS にコピーするには、Amazon S3 に保存された JSON ファイルまたはローカルファイルシステムのファイル (この例では *myStep.json*) に次の内容を配置します。 *j-3GYXXXXXX9I0K* をクラスター ID に置き換え、 *mybucket* を Amazon S3 バケット名に置き換えます。

```
[
  {
    "Name": "S3DistCp step",
    "Args": ["s3-dist-cp", "--s3Endpoint=s3.amazonaws.com", "--src=s3://mybucket/logs/j-3GYXXXXXX9I0J/node/", "--dest=hdfs:///output", "--srcPattern=.*daemons.*-hadoop-.*"],
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR",
    "Jar": "command-runner.jar"
  }
]
```

## S3DistCp ジョブが失敗した後のクリーンアップ

S3DistCp が指定されたファイルの一部または全部をコピーできない場合、コマンドまたはクラスターステップは失敗し、ゼロ以外のエラーコードを返します。この場合、S3DistCp は部分的にコピーされたファイルをクリーンアップしません。手動で削除する必要があります。

部分的にコピーされたファイルは、S3DistCp job の一意の識別子を使用して、サブディレクトリの HDFS tmp ディレクトリに保存されます。この ID はジョブの標準出力に含まれます。

例えば、ID が の S3DistCp ジョブの場合 *4b1c37bb-91af-4391-aaf8-46a6067085a6*、クラスターのマスターノードに接続し、次のコマンドを実行して、ジョブに関連付けられた出力ファイルを表示できます。

```
hdfs dfs -ls /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6/output
```

このコマンドによって以下のようなファイルのリストが返されます。

```
Found 8 items
-rw-r--r--  1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6/output/_SUCCESS
-rw-r--r--  1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6/output/part-r-00000
-rw-r--r--  1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6/output/part-r-00001
```

```
-rw-r--r--    1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-  
aaf8-46a6067085a6/output/part-r-00002  
-rw-r--r--    1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-  
aaf8-46a6067085a6/output/part-r-00003  
-rw-r--r--    1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-  
aaf8-46a6067085a6/output/part-r-00004  
-rw-r--r--    1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-  
aaf8-46a6067085a6/output/part-r-00005  
-rw-r--r--    1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-  
aaf8-46a6067085a6/output/part-r-00006
```

その後、以下のコマンドを実行して、ディレクトリとそのすべての内容を削除できます。

```
hdfs dfs rm -rf /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6
```



# Amazon EMR クラスターでのコマンドとスクリプトの実行

このトピックでは、クラスターでステップとしてコマンドまたはスクリプトを実行する方法について説明します。コマンドまたはスクリプトをステップとして実行することは、[クラスターに作業を送信する](#) 多くの方法の 1 つであり、次のような場合に役立ちます。

- Amazon EMR クラスターへの SSH アクセスがないとき
- bash コマンドまたはシェルスクリプトを実行してクラスターのトラブルシューティングを行うとき

スクリプトは、クラスターを作成するときにも、クラスターが WAITING 状態のときにも実行できません。ステップの処理が始まる前にスクリプトを実行するには、代わりにブートストラップアクションを使用します。ブートストラップアクションの詳細については、Amazon EMR 管理ガイドの[ブートストラップアクションを作成して追加のソフトウェアをインストールする](#) を参照してください。

Amazon EMR には、スクリプト、コマンド、およびその他のクラスター上のプログラムを実行するのに役立つ次のツールが用意されています。どちらのツールも、Amazon EMR 管理コンソールまたは AWS CLI を使用して起動できます。

## command-runner.jar

クラスターの Amazon EMR AMI にあります。command-runner.jar を使用して、クラスター上でコマンドを実行できます。command-runner.jar は、フルパスを使用せずに指定します。

## script-runner.jar

Amazon S3 (s3://<region>.elasticmapreduce/libs/script-runner/script-runner.jar) でホストされます。ここで、<region> は Amazon EMR クラスターが存在するリージョンです。script-runner.jar を使用して、ローカルまたはクラスター上の Amazon S3 に保存されたスクリプトを実行できます。ステップを送信するときに、script-runner.jar の完全な URI を指定する必要があります。

## スクリプトまたはコマンドを実行するためのカスタム JAR ステップの送信

Amazon EMR での command-runner.jar と script-runner.jar の一般的なユースケースをいくつか、次の AWS CLI 例に示します。

## Example : **command-runner.jar** を使用したクラスターでのコマンドの実行

`command-runner.jar` を使用するとき、ステップの引数のリストに、コマンド、オプション、および値を指定します。

次の AWS CLI 例では、`command-runner.jar` を起動する実行中のクラスターにステップを送信します。Args リストで指定されたコマンドは、`my-script.sh` というスクリプトを Amazon S3 から `hadoop` ユーザーのホームディレクトリにダウンロードします。次に、コマンドはスクリプトの権限を変更し、`my-script.sh` を実行します。

AWS CLI を使用するとき、Args リスト内の項目は、リスト要素の間に空白を含めずにコマンドで区切る必要があります。例えば、Args=[example-command, example-option, "example option value"] ではなく Args=[example-command,example-option,"example option value"] を使用します。

```
aws emr add-steps \  
--cluster-id j-2AXXXXXXGAPLF \  
--steps Type=CUSTOM_JAR,Name="Download a script from S3, change its permissions, and run it",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[bash,-c,"aws s3 cp s3://EXAMPLE-DOC-BUCKET/my-script.sh /home/hadoop; chmod u+x /home/hadoop/my-script.sh; cd /home/hadoop; ./my-script.sh"]
```

## Example : **script-runner.jar** を使用したクラスターでのスクリプトの実行

`script-runner.jar` を使用するとき、ステップの引数のリストで実行するスクリプトを指定します。

次の AWS CLI 例では、`script-runner.jar` を起動する実行中のクラスターにステップを送信します。この場合、`my-script.sh` という名前のスクリプトが Amazon S3 に保存されます。クラスターのマスターノードに保存されているローカルスクリプトを指定することもできます。

```
aws emr add-steps \  
--cluster-id j-2AXXXXXXGAPLF \  
--steps Type=CUSTOM_JAR,Name="Run a script from S3 with script-runner.jar",ActionOnFailure=CONTINUE,Jar=s3://us-west-2.elasticmapreduce/libs/script-runner/script-runner.jar,Args=[s3://EXAMPLE-DOC-BUCKET/my-script.sh]
```

## command-runner.jar のその他の使用方法

command-runner.jar を使用して、spark-submit や hadoop-streaming などのツールがあるクラスターに作業を送信することもできます。command-runner.jar を使用してアプリケーションを起動するときは、SPARK、STREAMING、PIG などの値を使用する代わりに、ステップタイプとして CUSTOM\_JAR を指定します。ツールの可用性は、クラスターにインストールしたアプリケーションによって異なります。

次のコマンド例では、command-runner.jar を使用して、spark-submit を使用するステップを送信します。Args リストにはコマンドとして spark-submit が指定され、その後に Spark アプリケーション *my-app.py* の Amazon S3 URI が引数と値とともに続きます。

```
aws emr add-steps \  
--cluster-id j-2AXXXXXXGAPLF \  
--steps Type=CUSTOM_JAR,Name="Run spark-submit using command-  
runner.jar",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[spark-submit,S3://  
DOC-EXAMPLE-BUCKET/my-app.py,ArgName1,ArgValue1,ArgName2,ArgValue2]
```

次の表に、command-runner.jar を使用して実行できるその他のツールを示します。

ツール名	説明
hadoop-streaming	Hadoop ストリーミングプログラムを送信します。コンソールと SDK では、これはストリーミングステップです。
hive-script	Hive スクリプトを実行します。コンソールと SDK では、これは、Hive のステップです。
pig-script	Pig スクリプトを実行します。コンソールと SDK では、これは、Pig のステップです。
spark-submit	Spark アプリケーションを実行します。コンソールでは、これは Spark のステップです。
hadoop-lzo	ディレクトリで <a href="#">Hadoop LZO インデクサ</a> を実行します。

ツール名	説明
s3-dist-cp	大量のデータを Amazon S3 から HDFS に分散コピーします。詳細については、「 <a href="#">S3DistCp (s3-dist-cp)</a> 」を参照してください。

# AWS 用語集

AWS の最新の用語については、「AWS の用語集リファレンス」の「[AWS 用語集](#)」を参照してください。

翻訳は機械翻訳により提供されています。提供された翻訳内容と英語版の間で齟齬、不一致または矛盾がある場合、英語版が優先します。