



Amazon RDS for My SQL、Amazon RDS for MariaDB、および Aurora My SQL互換 のデータのアーカイブ

AWS 規範ガイドンス



AWS 規範ガイド: Amazon RDS for My SQL、Amazon RDS for MariaDB、および Aurora My SQL互換 のデータのアーカイブ

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

Amazon の商標およびトレードドレスは、Amazon 以外の製品およびサービスに使用することはできません。また、お客様に誤解を与える可能性がある形式で、または Amazon の信用を損なう形式で使用することもできません。Amazon が所有していないその他のすべての商標は、Amazon との提携、関連、支援関係の有無にかかわらず、それら該当する所有者の資産です。

Table of Contents

序章	1
概要	1
目標とする成果	2
分割テーブルからのアーカイブ	4
分割されていないテーブルからのアーカイブ	6
Amazon S3 にデータを移動する	8
S3 SELECT INTO OUTFILE を使用する	8
データパイプラインを使用する	9
Glue AWS を使用する	9
アーカイブされたデータへのアクセス	13
標準ストレージクラス	13
S3 グレイシャーストレージクラス	14
ベストプラクティス	16
クリーンアップ	17
リソース	18
付録 I	19
付録 II	21
ドキュメント履歴	24
用語集	25
#	25
A	26
B	29
C	31
D	34
E	38
F	40
G	41
H	42
I	43
L	46
M	47
O	51
P	53
Q	56

R	56
S	59
T	63
U	64
V	65
W	65
Z	66
.....	lxvii

MySQL 用 Amazon RDS、MariaDB 用 Amazon RDS、および MySQL 互換の Aurora にデータをアーカイブする

Shyam Sunder Rakhecha、Abhishek Karmakar、Oliver Francis、Saumya Singh アマゾンウェブサービス (AWS)

2023 年 4 月([ドキュメント履歴](#))

履歴データをアーカイブする必要性は、さまざまなユースケースに起因する可能性があります。アプリケーションがアーカイブ機能なしで設計されている場合があり、時間の経過とともにビジネスが拡大すると、大量の履歴データが生成される可能性があります。これは必然的にパフォーマンスの低下につながります。また、組織内のコンプライアンス要件により、履歴データを保持する場合があります。

このガイドでは、アプリケーションへの影響を最小限に抑えながら Amazon Simple Storage Service (Amazon S3) に履歴データをアーカイブし、必要なときにアーカイブされた情報を取得する方法について説明します。

概要

このガイドでは、MySQL 用の Amazon リレーショナルデータベースサービス (Amazon RDS)、MariaDB 用 Amazon RDS、およびアマゾンウェブサービス (AWS) クラウドの Amazon Aurora MySQL 互換エディションの大きなテーブルから履歴データをアーカイブするさまざまな方法について説明します。このガイドでは、分割されたテーブルのデータと、分割されていない大きなテーブルにあるデータの両方をアーカイブする方法を学びます。ガイドに示されているアプローチを実装すると、重要な履歴データを保存して今後の分析に備えながら、ライブデータのサイズを小さくすることができます。

テーブルデータを定期的にアーカイブすると、テーブル内のライブデータセットの量が少なくなり、読み取りと書き込みが速くなり、アプリケーションのパフォーマンスが向上します。定期的なデータアーカイブは、「オペレーショナル・エクセレンス」と「パフォーマンス効率」の柱に該当します [よく設計されたフレームワーク](#)。古いデータを Amazon シンプルストレージサービス (Amazon S3) に移動し、Amazon RDS インスタンスまたは Aurora MySQL 互換クラスターでアーカイブされたデータをクリーンアップすると、ストレージコストを節約できます。これはコスト最適化の柱に合致し、AWS での不必要なコストを回避するのに役立ちます。

ターゲットを絞ったビジネス成果

このガイドでは、以下のビジネス成果に焦点を当てています。

- ユーザーエクスペリエンスの向上
- データコンプライアンス要件が満たされました
- ストレージコストの削減
- 整理されたデータ

ユーザーエクスペリエンスの向上

履歴データを保持するデータベースは、テーブルとインデックスが大きいため、パフォーマンスが低下する可能性があります。履歴データをアーカイブすると、テーブルとインデックスがスリムになります。これは、データベースとやり取りする顧客向けの API 操作に直接的なプラスの影響を与えます。

データコンプライアンス要件が満たされました

金融サービス、公共機関、医療などの業界では、厳しいアーカイブ要件があります。MySQL 用 Amazon RDS、MariaDB 用 Amazon RDS、または Amazon S3 の Aurora MySQL 互換データベースにあるアプリケーションデータをアーカイブすることで、次のような規制コンプライアンスの要件を満たすことができます。

- Payment Card Industry Data Security Standard (PCI DSS)
- 医療保険の相互運用性と説明責任に関する法律 (HIPAA) および経済的および臨床的健康のための医療情報技術 (HITECH) 法
- Federal Risk and Authorization Management Program (FedRAMP)
- 一般データ保護規則 (GDPR)
- 連邦情報処理規格 (FIPS) 140-2
- 米国国立標準技術研究所 (NIST) 800—171

ストレージコストの削減

Amazon RDS にデータを保存すると、ストレージコストが高くなり、IOPS も高くなります。1 か月あたり GB あたりのストレージコストを比較すると、[MySQL 用アマゾン RDS マルチ AZ GP2](#)のそ

れと[アマゾン S3 グレイシャー](#)にus-east-1AWS リージョンでは、S3 Glacier のストレージコストは Amazon RDS のストレージコストよりも約 57 倍低くなっています。

整理されたデータ

アプリケーションが頻繁にアクセスする有益なデータをデータベースに保存しておくといでしょう。ただし、アプリケーションは大量のデータを生成しますが、そのデータはそれほど頻繁には必要ないか、古くなります。これらの記録はアーカイブして所定の場所に保存できるため、費用対効果が高く、アプリケーションのパフォーマンスに影響を与えません。

分割テーブルへのデータのアーカイブ

MySQL はサポートしています [分割](#)InnoDB ストレージエンジン用で、この機能を使用して大きなテーブルを分割できます。テーブル内のパーティションは個別の物理テーブルとして格納されますが、分割されたテーブルを操作する SQL はテーブル全体を読み取ります。これにより、実行せずにテーブルから不要なパーティションを自由に削除できます。row-by-rowを削除して、データベース内の履歴行をアーカイブできるようにします。

次のコード例を考えてみましょう。TABLE orders内に存在するorderprocessingスキーマ。その履歴データはパーティションにありますphistoricalには、2021年以前のデータが含まれています。同じ表では、アプリケーションレベルのホットデータが、2022年の各月のライブパーティションにあります。パーティション内のデータをアーカイブするにはphistorical、アーカイブを作成できますtable orders_2021_and_older同じ構造でarchiveスキーマ。その後、MySQLを使用できます。 [パーティションの交換](#)パーティションを移動するにはphistoricalそのテーブルにアーカイブテーブルは分割されていないことに注意してください。アーカイブ後、データを検証してAmazon S3に移動できます。

```
CREATE TABLE orders (
  orderid bigint NOT NULL AUTO_INCREMENT,
  customerid bigint DEFAULT NULL,
  .....
  .....
  order_date date NOT NULL,
  PRIMARY KEY (`orderid`, `order_date`))
PARTITION BY RANGE (TO_DAYS(order_date)) (
  PARTITION pstart VALUES LESS THAN (0),
  PARTITION phistorical VALUES LESS THAN (TO_DAYS('2022-01-01')),
  PARTITION p2022JAN VALUES LESS THAN (TO_DAYS('2022-02-01')),
  PARTITION p2022FEB VALUES LESS THAN (TO_DAYS('2022-03-01')),
  PARTITION p2022MAR VALUES LESS THAN (TO_DAYS('2022-04-01')),
  PARTITION p2022APR VALUES LESS THAN (TO_DAYS('2022-05-01')),
  PARTITION p2022MAY VALUES LESS THAN (TO_DAYS('2022-06-01')),
  PARTITION p2022JUN VALUES LESS THAN (TO_DAYS('2022-07-01')),
  PARTITION p2022JUL VALUES LESS THAN (TO_DAYS('2022-08-01')),
  PARTITION p2022AUG VALUES LESS THAN (TO_DAYS('2022-09-01')),
  PARTITION p2022SEP VALUES LESS THAN (TO_DAYS('2022-10-01')),
  PARTITION p2022OCT VALUES LESS THAN (TO_DAYS('2022-11-01')),
  PARTITION p2022NOV VALUES LESS THAN (TO_DAYS('2022-12-01')),
  PARTITION p2022DEC VALUES LESS THAN (TO_DAYS('2023-01-01')),
  PARTITION pfuture VALUES LESS THAN MAXVALUE
```

```
);  
CREATE TABLE orders_2021_and_older (  
    orderid bigint NOT NULL AUTO_INCREMENT,  
    customerid bigint DEFAULT NULL,  
    .....  
    .....  
    order_date date NOT NULL,  
    PRIMARY KEY (`orderid`,`order_date`));  
mysql> alter table orderprocessing.orders exchange partition phistorical with table  
archive.orders_2021_and_older;  
Query OK, 0 rows affected (0.33 sec)
```

使用する場合EXCHANGE PARTITION履歴データをアーカイブする機能については、以下のベストプラクティスをお勧めします。

- アプリケーションにアーカイブデータを保存するための別のスキーマを作成します。このスキーマには、アーカイブされたデータを格納するアーカイブテーブルが含まれます。アーカイブスキーマのアーカイブテーブルは、インデックスと主キーを含め、ライブテーブルと同じ構造でなければなりません。ただし、ターゲットアーカイブテーブルを分割テーブルにすることはできません。MySQLでは、2つのパーティションテーブル間でパーティションを交換することは許可されていません。
- アーカイブテーブルには、保存されている履歴データを識別しやすい命名規則に従ってください。これは、このデータを Amazon S3 に移動する監査タスクや設計ジョブを実行する場合に便利です。
- を実行してくださいEXCHANGE PARTITIONAurora MySQL 互換ライター、MySQL 用 Amazon RDS、または MariaDB 用の Amazon RDS インスタンスにトラフィックが入らない場合に、ダウンタイムウィンドウに表示されるデータ定義言語 (DDL) ステートメント。

実行できるかもしれませんがEXCHANGE PARTITIONアプリケーションまたはマイクロサービスのトラフィックの少ない時間帯に ただし、パーティションテーブルには書き込みや選択は行わないか、ごくわずかであるはずで、実行時間の長い既存の select クエリが原因でEXCHANGE PARTITIONDDL が待機しているため、データベース上でリソース競合が発生しています。実行前にこれらの条件がすべて満たされていることを確認するデザインスクリプトEXCHANGE PARTITIONお使いのシステム上。

アプリケーションの設計が分割データに対応していて、現在分割されていないテーブルがある場合は、データのアーカイブをサポートするためにデータを分割テーブルに移動することを検討してください。詳細については、[MySQL ドキュメント](#)を参照してください。

分割されていないテーブルからのデータのアーカイブ

パーティショニングが不可能なデータベーステーブルでは、Percona Toolkit を使用できます。[pt-アーカイバー](#) テーブルのデータを MySQL データベースの別のテーブルにアーカイブするツール。

pt-archiver ツールを使用して、大きなテーブルのレコードを他のテーブルやファイルにアーカイブします。これは読み取り/書き込みツールです。つまり、データをアーカイブした後にソーステーブルからデータを削除するので、ソースデータの削除を個別に管理する必要はありません。このスクリプトの主な目的は、既存のオンライントランザクション処理 (OLTP) クエリロード (付録 I を参照) に影響を与えずにテーブルから古いデータをアーカイブし、そのデータを同じサーバーまたは別のサーバー上の別のテーブルに挿入することです。

ダウンロードできます [ペルコナツールキット](#) として [インストール](#) ローカルマシン、またはデータベースへの接続元となる Amazon Elastic Compute Cloud (Amazon EC2) インスタンス上にあります。pt-archiver ツールを実行するには、次の構文を使用します。

```
pt-archiver --source h=<HOST>,D=<DATABASE>,t=<TABLE>,u=<USER>,p=<PASSWORD> --dest  
h=<HOST>,D=<DATABASE>,t=<TABLE> --where "'1=1'" --statistics
```

交換してください HOST、DATABASE、TABLE、および USER ソースとターゲットのデータベースの詳細と認証情報が必要です。

使用することもできます [AWS バッチ](#) このジョブをテーブル用に作成してスケジュールできます。

pt-archiver ツールを使用してテーブルのデータをアーカイブする場合は、次の点を考慮してください。

- ソーステーブルに主キーがあると、このツールのパフォーマンスが向上します。テーブルに主キーがない場合は、[ユニークな列にインデックスを作成](#) これは pt-archiver がテーブルのすべての行を調べてアーカイブするのに役立ちます。
- デフォルトでは、pt-archiver はテーブルをアーカイブした後にデータを削除します。本番サーバーで実行する前に、必ず次の方法でアーカイブジョブをテストしてください。--dry-run。または、--no-delete オプション。
- pt-archiver ツールは、システムの負荷に基づいてアーカイブの速度を調整します (付録 II を参照)。負荷が高いと、アーカイブのパフォーマンスが低下することが予想されます。

pt-archiver を実行すると、アーカイブされたデータはアーカイブスキーマの対応するテーブルにあるはずですが、そこから Amazon S3 に移動できます。

アーカイブされたテーブルデータを Amazon S3 に移動する

Amazon Simple Storage Service (Amazon S3) は、アーカイブされたデータの自然なターゲットです。99.999999999% の耐久性を提供し、データベースストレージよりも安価です。

さらに、Amazon S3 には、取得パターンに基づいて料金が設定されたストレージクラスが組み込まれています。データの取得頻度に基づいて、オフロードされた S3 オブジェクトを低価格のストレージ階層に移行するオプションがあります。ストレージクラスと料金の詳細については、[Amazon S3 ドキュメント](#)」を参照してください。

MySQL インスタンスのフリートを使用するアプリケーションの場合、Amazon S3 にオフロードすると、次の条件を満たすデータに対するコストを節約できます。

- 効率を向上させるには、データベースからアーカイブする必要があります
- すぐには必要ない、またはビジネスプロセスにほとんど必要ない
- 監査要件のため、長期間保持する必要があります

MySQL データは、次の方法でアーカイブできます。

- を使用してデータをエクスポートします `SELECT INTO OUTFILE S3`。
- Data Pipeline を使用して AWS データをエクスポートします。
- Glue AWS を使用してデータをエクスポートします。

S3 SELECT INTO OUTFILE を使用してデータをエクスポートする

Aurora My SQL互換 DB から Amazon S3 にデータを直接コピーするには、ステートメントを使用できます `SELECT INTO OUTFILE S3`。この SQL ステートメントはテーブルで実行でき、必要な行数は、最大サイズが 6 GB のカンマ区切り値 (CSV) ファイルとしてオフロードできます。6 GB のしきい値を超えると、複数の .csv ファイルが作成されます。

エクスポートを機能させるには、以下を設定します。

- AWS Identity and Access Management (IAM) のロールとポリシー
- コマンドを実行するためにユーザーに付与されたデータベース許可
- オフロード用の Amazon S3 の場所

詳細については、「[Amazon Aurora ドキュメント](#)」を参照してください。

注：S3 オブジェクトをコスト削減ストレージ階層に移行するには、データをエクスポートする [S3 バケットで Amazon S3 ライフサイクルルール](#)を設定することをお勧めします。S3

Data Pipeline を使用してAWSデータをエクスポートする

AWS Data Pipeline は、データ駆動型ワークフロー用のウェブサービスです。Data Pipeline を使用して、Amazon RDS for MySQL から Amazon S3 へのデータの移動を自動化できます。Data Pipeline は一連のアクションまたはタスクを連続的に実行し、データを .csv ファイルとして Amazon S3 にエクスポートします。

パイプラインは、次のいずれかの方法で作成できます。

- コンソールで、ソースのテンプレートRDS「MySQLテーブルの S3 へのフルコピー」を選択し、パラメータの値を指定します。
- からAWSCLI、データパイプラインオブジェクトをとして定義しJSON、パイプラインの作成、アップロード、アクティブ化を行います。

Data Pipeline を使用する利点は、エクスポートジョブをスケジュールしてモニタリングできることです。さらに、大きなテーブルでは、Amazon EMRの代わりに Amazon を使用できますEC2。詳細と例については、[AWS「Data Pipeline ドキュメント」](#)を参照してください。

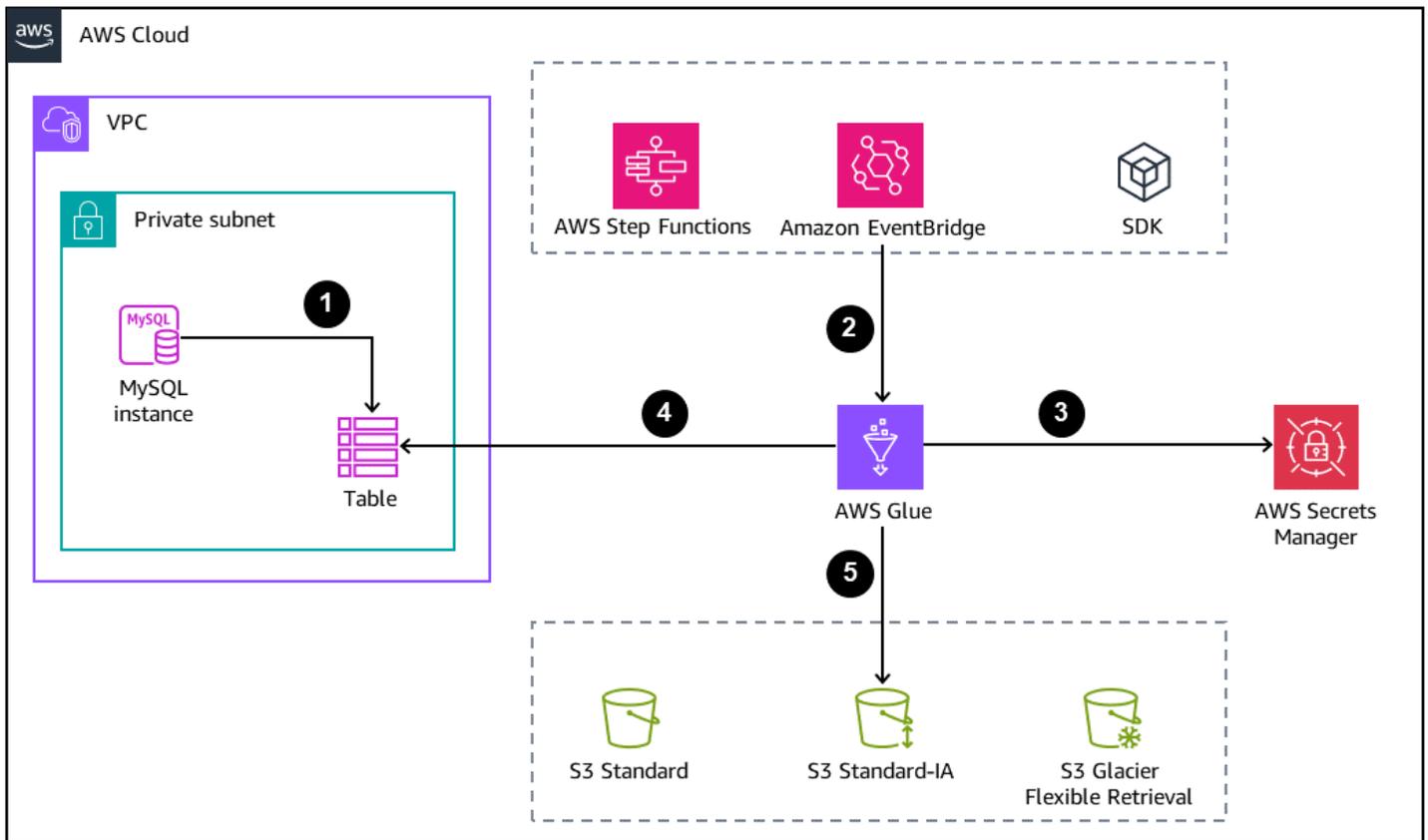
注：S3 オブジェクトをコスト削減ストレージ階層に移行するには、データエクスポートが実行されている [S3 バケットで Amazon S3 ライフサイクルルール](#)を設定することをお勧めします。S3

Glue AWS を使用してデータをエクスポートする

ビッグデータシナリオ用のサーバーレス分析サービスである Glue AWS を使用して、Amazon S3 にデータをアーカイブSQLできます。AWS Glue は、多くのデータベースソースをサポートする広く使用されている分散クラスターコンピューティングフレームワークである Apache Spark を使用しています。

データベースから Amazon S3 へのアーカイブデータのオフロードは、Glue AWS ジョブ内の数行のコードで実行できます。AWS Glue が提供する最大の利点は、水平スケーラビリティと pay-as-you-go モデルであり、運用効率とコスト最適化を提供します。

次の図は、データベースアーカイブの基本的なアーキテクチャを示しています。



1. データベースSQLは、Amazon S3 にオフロードされるアーカイブテーブルまたはバックアップテーブルを作成します。
2. Glue AWS ジョブは、次のいずれかの方法で開始されます。
 - [AWS Step Functions](#) ステートマシン内のステップとして同期する
 - [Amazon EventBridge](#) イベントによる非同期
 - AWS CLI または を使用した手動リクエストによる [AWS SDK](#)
3. DB 認証情報は [AWS Secrets Manager](#) から取得されます。
4. AWS Glue ジョブは、Java Database Connectivity (JDBC) 接続を使用してデータベースにアクセスし、テーブルを読み込みます。
5. AWS Glue は Amazon S3 でデータを Parquet 形式で書き込みます。Parquet 形式は、オープンで列指向、省スペースのデータ形式です。

Glue AWS ジョブの設定

AWS Glue ジョブを意図したとおりに動作させるには、以下のコンポーネントと設定が必要です。

- [AWS Glue 接続](#) — これは、データベースにアクセスするためにジョブにアタッチする Glue Data Catalog オブジェクトです。AWSジョブは、複数のデータベースを呼び出すための複数の接続を持つことができます。接続には、安全に保存されたデータベース認証情報が含まれています。
- [GlueContext](#) – これは、GlueContext クラス上のカスタムラッパーであり、Amazon [SparkContext](#) S3 およびデータベースソースとやり取りするための高次のAPIオペレーションを提供します。Amazon S3 これにより、Data Catalog との統合が可能になります。また、Glue 接続内で処理されるデータベース接続のドライバーに依存する必要もなくなります。さらに、GlueContext クラスは Amazon S3 APIオペレーションを処理する方法を提供します。これは元の SparkContext クラスでは不可能です。
- IAM ポリシーとロール – Glue AWS は他の AWS サービスとやり取りするため、必要な最小限の特権で適切なロールを設定する必要があります。Glue AWS とやり取りするために適切なアクセス許可を必要とするサービスには、次のようなものがあります。
 - Amazon S3
 - AWS Secrets Manager
 - AWS Key Management Service (AWS KMS)

ベストプラクティス

- オフロードする行が多数あるテーブル全体を読み取る場合は、リードレプリカエンドポイントを使用して、メインライターインスタンスのパフォーマンスを低下させることなく読み取りスループットを向上させることをお勧めします。
- ジョブの処理に使用されるノード数の効率を高めるには、Glue AWS 3.0 で[自動スケーリング](#)を有効にします。
- S3 バケットがデータレイクアーキテクチャの一部である場合は、データを物理パーティションに整理してオフロードすることをお勧めします。パーティションスキームはアクセスパターンに基づいている必要があります。日付値に基づくパーティション化は、最も推奨されるプラクティスの一つです。
- Parquet や Optimized Row Columnar (ORC) などのオープンな形式にデータを保存すると、Amazon Athena や Amazon Redshift などの他の分析サービスでデータを使用できるようになります。
- オフロードされたデータを他の分散サービスで読み取り最適化するには、出力ファイルの数を制御する必要があります。小さなファイルが多いのではなく、大きなファイルの数が少ないことがほとんどの場合有益です。Spark には、パートファイルの生成を制御するための設定ファイルとメソッドが組み込まれています。

- 定義上アーカイブされたデータは、多くの場合、アクセスされるデータセットです。ストレージのコスト効率を実現するには、Amazon S3 クラスをより安価な階層に移行する必要があります。これは、次の2つの方法を使用して実行できます。
- オフロード中に階層を同期的に移行する – オフロードされたデータをプロセスの一部として移行する必要があることが事前にわかっている場合は、Amazon S3 にデータを書き込むのと同じ AWS Glue ジョブ内で GlueContext メカニズム `transition_s3_path` を使用できます。
https://docs.aws.amazon.com/glue/latest/dg/aws-glue-api-crawler-pyspark-extensions-glue-context.html#aws-glue-api-crawler-pyspark-extensions-glue-context-transition_s3_path Amazon S3
- [S3 ライフサイクル](#)を使用した非同期移行 — Amazon S3 ストレージクラスの移行と有効期限の適切なパラメータを使用して S3 ライフサイクルルールを設定します。Amazon S3 これをバケットに設定した後は、永続化されます。
- データベースがデプロイされている仮想プライベートクラウド (VPC) 内で、[十分な IP アドレス範囲](#)を持つサブネットを作成して設定します。これにより、大量のデータ処理ユニット (DPUs) AWS が設定されている場合にネットワークアドレスの数が不足することによる Glue ジョブの失敗を回避できます。

Amazon S3 のアーカイブデータへのアクセス

Amazon S3 には、データの内容を読み取るためのツールが多数用意されています。ただし、ストレージクラスによっては、いくつかの前処理手順が必要になる場合があります。このセクションには次の内容が含まれます。

- AWS Glue または Amazon S3 Select を使用して、アーカイブされた S3 オブジェクトを標準ストレージクラスで読み取る
- Amazon S3 Glacier Select または S3 バッチオペレーションを使用して S3 Glacier ストレージクラスでアーカイブされた S3 オブジェクトを読み取る
- ベストプラクティス

Standard ストレージクラスでアーカイブされた S3 オブジェクトを読み取る

Amazon S3 ストレージクラスでアーカイブされた S3 オブジェクトは、AWS Glue または Amazon S3 Select のいずれかを使用して読み取ることができます。

AWS グループの使用

MySQL から Amazon S3 にオフロードされたデータは、リレーショナルデータベース管理システム (RDBMS) と同じ構造の剛性と一貫性を維持します。

[AWS グループクローラー](#) S3 オブジェクトをクローリングしてデータ型を推測し、テーブルメタデータを外部テーブル DDL として作成します。クローラージョブを設定するときは、Amazon S3 をソースとして使用し、すべてのデータファイルが作成される S3 プレフィックスの場所を指定します。構成には、以下を含めてください。

- クローラー実行オプション
- オプションのテーブルプレフィックスプリファレンス
- テーブルを作成するためのターゲットデータベース
- 必要な権限を持つ IAM ロール

ジョブを呼び出すと、データをスキャンしてスキーマを推測し、保存します。[AWS グループデータカタログ](#)なので[AWS グループテーブル](#)。AWS Glue テーブルは基本的に外部テーブルであり、次のよう

な分析サービスを使用して通常のデータベーステーブルと同様に SQL ステートメントでクエリできます。[アマゾンアテナ](#)、[アマゾン赤方偏移スペクトル](#)、そしてアパッチハイブオン[アマゾン EMR](#)。クローラの詳細については、[を参照してください](#)。[AWS グループのドキュメンテーション](#)。

列ヘッダーが指定されている.csv ファイルの場合、生成されるテーブルの列名には同じフィールド名が反映されます。データ型は、データオブジェクトの値に基づいて推測されます。

Parquet ファイルの場合、スキーマはデータ自体に保存され、生成されるテーブルには同じフィールド名とデータ型が反映されます。

または、Athena 内で DDL を手動で実行して、必要な列名とデータ型でテーブル定義を作成することもできます。これにより、データカタログ内にテーブル定義が作成されます。Athena テーブルの作成の詳細については、[を参照してください](#)。[Amazon Athena ドキュメンテーション](#)。

注記:CSV ファイルにヘッダー行がない場合、クローラはフィールド名を汎用 c_0、c_1、c_2、... として作成します。

Amazon S3 セレクトを使用する

Amazon S3 Select を使用すると、SQL 式を使用して S3 オブジェクトをプログラマ的に読み取ることができます。API オペレーションは AWS CLI コマンドを使用して呼び出すことができます。select-object-content または、Boto3 などの SDK を使用して操作を呼び出す select_object_content パイソンから。

API オペレーションは SQL ステートメントをパラメーターとしてサポートし、JSON タイプと Parquet タイプのファイルのみを読み取ることができます。出力は出力ファイルとしてリダイレクトできます。

これらの操作は、S3 オブジェクトごとに呼び出されます。複数のファイルの場合は、操作を再帰的に実行します。

AWS CLI を使用して操作を実行する方法の詳細については、[を参照してください](#)。[AWS CLI ドキュメンテーション](#)。Python SDK Boto3 を使用して S3 セレクトを実行する方法の詳細については、[以下を参照してください](#)。[Boto3 ドキュメンテーション](#)。

S3 Glacier ストレージクラスによるアーカイブされた S3 オブジェクトの読み取り

Amazon S3 Glacier クラスは特別なストレージクラスで、価格は安価ですが検索に時間がかかります。S3 標準オブジェクトとは異なり、S3 Glacier オブジェクトは AWS Glue テーブルとして読み取

ることができません。データを分析クエリやレポートに使用できるようにするには、まず S3 Glacier オブジェクトを復元します。復元は時間をかけて行われる非同期プロセスで、保存期間があります。オブジェクトが復元されたら、S3 Standard オブジェクトとして別の場所にコピーできます。保持期間を過ぎると、復元されたオブジェクトは Amazon S3 Glacier に戻ります。

Amazon S3 グレイシャーセレクトの使用

Amazon S3 Select を S3 スタンダードで使用するのと同様に、S3 Glacier オブジェクトにクエリを実行してデータのサブセットを取得できます。これにより、オブジェクトの復元やその他の AWS 分析サービスなどの前処理を必要とせずに、プログラムでデータにアクセスできます。例については、[を参照してください](#)。[リクエスト構文](#) Boto3 を使用する場合 `initiate_job` S3 Glacier Select データを読み取る操作。

S3 バッチオペレーションの使用

S3 バッチオペレーションでは、エクサバイトのデータを含む数十億ものオブジェクトにおよぶ大規模なバッチオペレーションを Amazon S3 で行うことができます。Amazon S3 は、進捗状況の追跡、通知の送信、すべてのアクションの詳細な完了レポートの保存を行い、フルマネージド型の監査可能なサーバーレスエクスペリエンスを提供します。

S3 バッチオペレーションは [復元](#) オペレーション。これにより、次のストレージ階層の S3 オブジェクトの復元が開始されます。

- S3 Glacier Flexible Retrieval または S3 Glacier Deep Archive ストレージクラスにアーカイブされたオブジェクト
- Archive アクセス階層または Deep Archive アクセス階層の S3 Intelligent-Tiering ストレージクラスを介してアーカイブされたオブジェクト

バッチオペレーションは、プログラムと Amazon S3 コンソールの両方で呼び出すことができます。入力には、復元するリストオブジェクトを含む.csv マニフェストファイルが必要です。

を使用できます [アマゾンS3インベントリ](#) バッチ作業の入力としてレポートを作成します。インベントリレポートはバケット用に設定され、特定のプレフィックスのオブジェクトに限定できます。これは自動レポートで、CSV、ORC、またはParquet形式で毎週または毎日生成されます。

インベントリレポートの設定の詳細については、[を参照してください](#)。[アマゾン S3 ドキュメンテーション](#)。Boto3 を使用して S3 バッチオペレーションジョブを作成する方法については、[を参照してください](#)。[Boto3 ドキュメンテーション](#)。

ベストプラクティス

アーカイブされたデータにアクセスするには、次のベストプラクティスをお勧めします。

- Amazon S3 Select と Amazon S3 Glacier Select は、データのサブセットが基本的な SQL 式を使用して出力される場合に適しています。選択したデータセットを導き出すためにプログラムで Amazon S3 にクエリする必要があるアプリケーションでは、パフォーマンスが大幅に向上し、多くの場合 400% も向上します。詳細については、ブログ記事を参照してください [S3 セレクトと S3 Glacier Select — オブジェクトのサブセットの取得](#)。
- S3 セレクトと S3 Glacier Select は、CSV、JSON、および Parquet 形式のデータの読み取りのみをサポートしています。ORC や Avro などの他のオープンフォーマットについては、Amazon Athena または Amazon Redshift を使用できます。
- 膨大なアーカイブデータセットの場合は、データの上に AWS Glue テーブルを作成して、Athena や Amazon Redshift などのクエリエンジンを使用して読み込めるようにすることをお勧めします。Athena と Amazon Redshift はどちらも、クエリのパフォーマンスを水平スケーリングできます。また、pay-per-query モデル。1 回限りのクエリシナリオでは費用対効果が高い。さらに、Amazon Redshift にはアドバンスドクエリアクセラレーター (AQUA) エンジンが組み込まれているため、追加コストなしで読み取りパフォーマンスが向上します。
- Amazon S3 に定期的にオフロードされるアーカイブデータは、ヒープダンプとして保存しないでください。代わりに、新しいパーティションとして保存する必要があります。日付パーティションは、データを日付ディメンションに分割します (たとえば、year=<value>/month=<value>/day=<value>)。これは次の 2 つの状況で非常に有益です。
 - AWS Glue テーブルが AWS Glue クローラーによって作成されている場合、これらのパーティションは疑似列として機能します。これにより、スキャンされるデータが範囲クエリのパーティションに制限されるため、読み取りパフォーマンスが向上します。
 - これは、オブジェクトのサブセットのみを S3 Standard として復元する場合の S3 Glacier 復元操作に役立ちます。
- AWS Glue クローラーは、Amazon S3 に保存されたアーカイブデータを物理的に分割すると大きな価値を発揮します。データが新しいプレフィックスパーティションとしてオフロードされるたびに、クローラは新しいパーティションのみをスキャンし、そのパーティションのメタデータを更新します。テーブルのスキーマが変更された場合、それらの変更はパーティションレベルのメタデータにキャプチャされます。

アーカイブテーブルのクリーンアップ

アーカイブプロセスの最終段階は、アーカイブスキーマのテーブルをクリーンアップすることです。これは、アーカイブデータが Amazon S3 に安全にアーカイブされていることを確認した後に行うことができます。アプリケーションへの影響を避けるため、スケジュールされたダウンタイムやメンテナンス期間、またはアプリケーションのトラフィックが非常に少ない時間帯にアーカイブスキーマテーブルを削除することをお勧めします。これらのテーブルはアプリケーションから積極的にクエリされるわけではないので、進行中のトランザクションへの影響について心配する必要はありません。それでも、ダウンタイム中に DDL を実行するのがベストプラクティスです。

大規模なアーカイブスキーマテーブルのストレージを解放すると、Amazon Aurora は [動的サイズ変更](#) ストレージコストを節約するのに役立ちます。

リソース

ドキュメント

- [Amazon Aurora](#)
- [AWS Data Pipeline](#)
- [AWS Glue](#)
- [Amazon S3](#)
- [Amazon Athena を使用して Amazon S3 インベントリをクエリする](#)
- [Amazon S3 インベントリの設定](#)
- [Amazon S3 オブジェクトに対する大規模なバッチオペレーションの実行](#)
- [Boto3 ドキュメント](#)
- [MySQL パーティショニングドキュメント](#)

料金表

- [MySQL 用 Amazon RDS マルチ AZ GP2 料金表](#)
- [Amazon S3 グレイシャーの価格](#)

Tools

- [ペルコナツールキット](#)
- [pt-アーカイバー](#)
- [システムベンチ](#)

ブログ投稿

- [S3 セレクトと S3 Glacier Select — オブジェクトのサブセットの取得](#)
- [pg_partman と Amazon S3 を使用して PostgreSQL 用の Amazon RDS と PostgreSQL 互換の Amazon Aurora のデータをアーカイブおよび消去する](#)

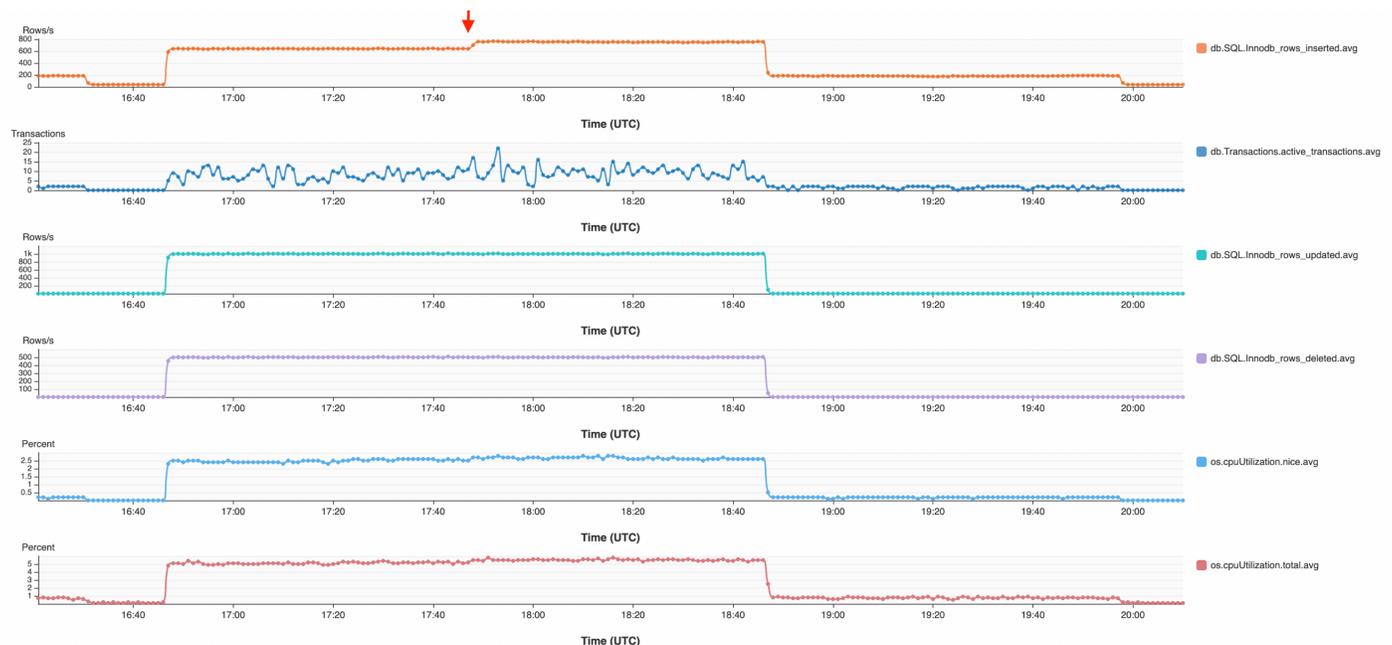
付録 I

すべてのテストは、上で実行されている Amazon RDS for MySQL インスタンスで実行されます。db.r6g.8xlargeインスタンスクラス。

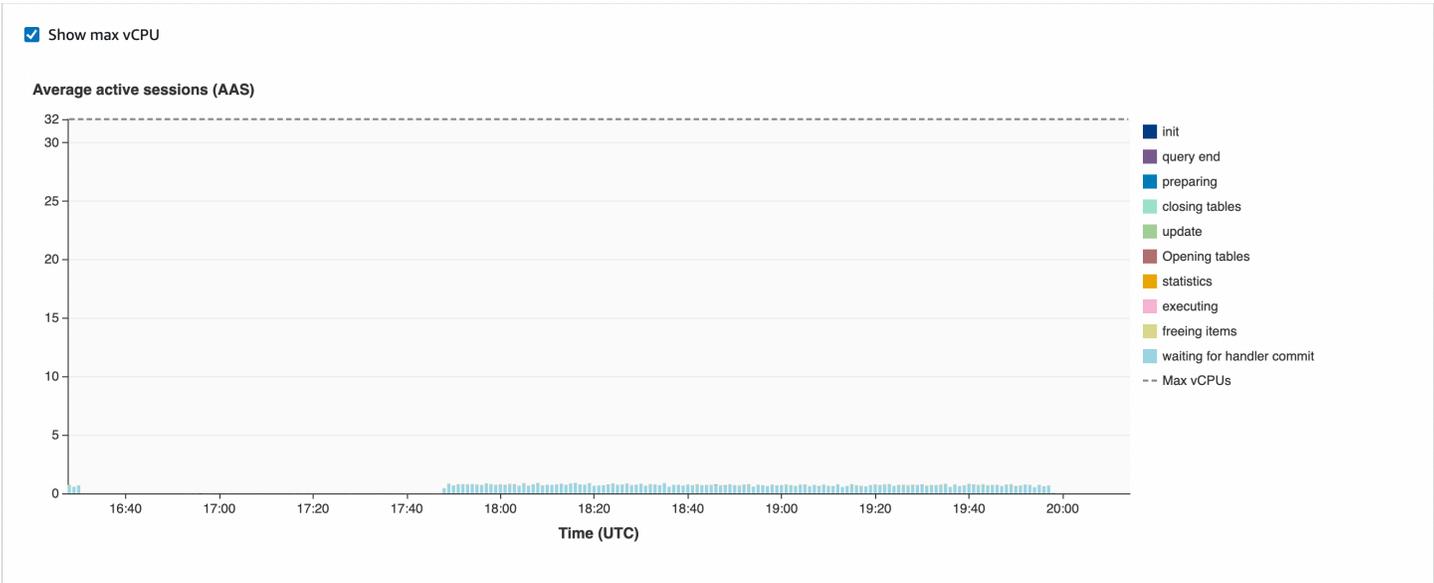
以下のシステムベンチコマンドを使用して、データベースへのロードの準備と実行を行いました。

```
sysbench oltp_read_write --db-driver=mysql --mysql-db=<DATABASE> --mysql-user=<USER> --mysql-password=<PASSWORD> --mysql-host=<ENDPOINT> --tables=500 --table-size=2000000 --threads=500 prepare
sysbench oltp_read_write --db-driver=mysql --mysql-db=employees --mysql-user=admin --mysql-password=qwertyuiop --mysql-host=mysql8.cbbhujzeoxed.us-east-1.rds.amazonaws.com --tables=500 --rate=500 --time=7200 run
```

次のグラフでは、OLTPワークロードが実行中で、矢印のマークが付いた場所でpt-archiverプロセスが開始されました。



pt-archiver を並行して実行しても CPU 使用率に大きな変化はありません。つまり、pt-archiver は実行中の OLTP クエリに影響を与えないことがわかります。



付録 II

このセクションでは、さまざまなシナリオでの pt-archiver ツールのベンチマーク結果を示します。ザ・[システムベンチ](#)このテストでは、ツールを使用してデータベースに負荷をかけます。すべてのテストは、以下で実行されている Amazon RDS for MySQL インスタンスで実行されます。db.r6g.8xlarge インスタンスクラス。

次の sysbench コマンドを使用して、データベースへのロードの準備と実行を行いました。

```
sysbench oltp_read_write --db-driver=mysql --mysql-db=<DATABASE> --mysql-user=<USER> --mysql-password=<PASSWORD> --mysql-host=<ENDPOINT> --tables=1000 --table-size=2000000 --threads=500 prepare
sysbench oltp_read_write --db-driver=mysql --mysql-db=<DATABASE> --mysql-user=<USER> --mysql-password=<PASSWORD> --mysql-host=<ENDPOINT> --tables=1000 --rate=500 --threads=500 run
```

主キーがなく、インデックスが 1 つしかないテーブルをアーカイブする (データベースに負荷をかけない)

```
Started at 2022-11-07T05:29:12, ended at 2022-11-07T06:03:31
Action      Count          Time           Pct
commit      600050         1715.3582     83.31
select      300025         166.5470      8.09
inserting   300024         165.4025      8.03
other       0              11.6644       0.57
```

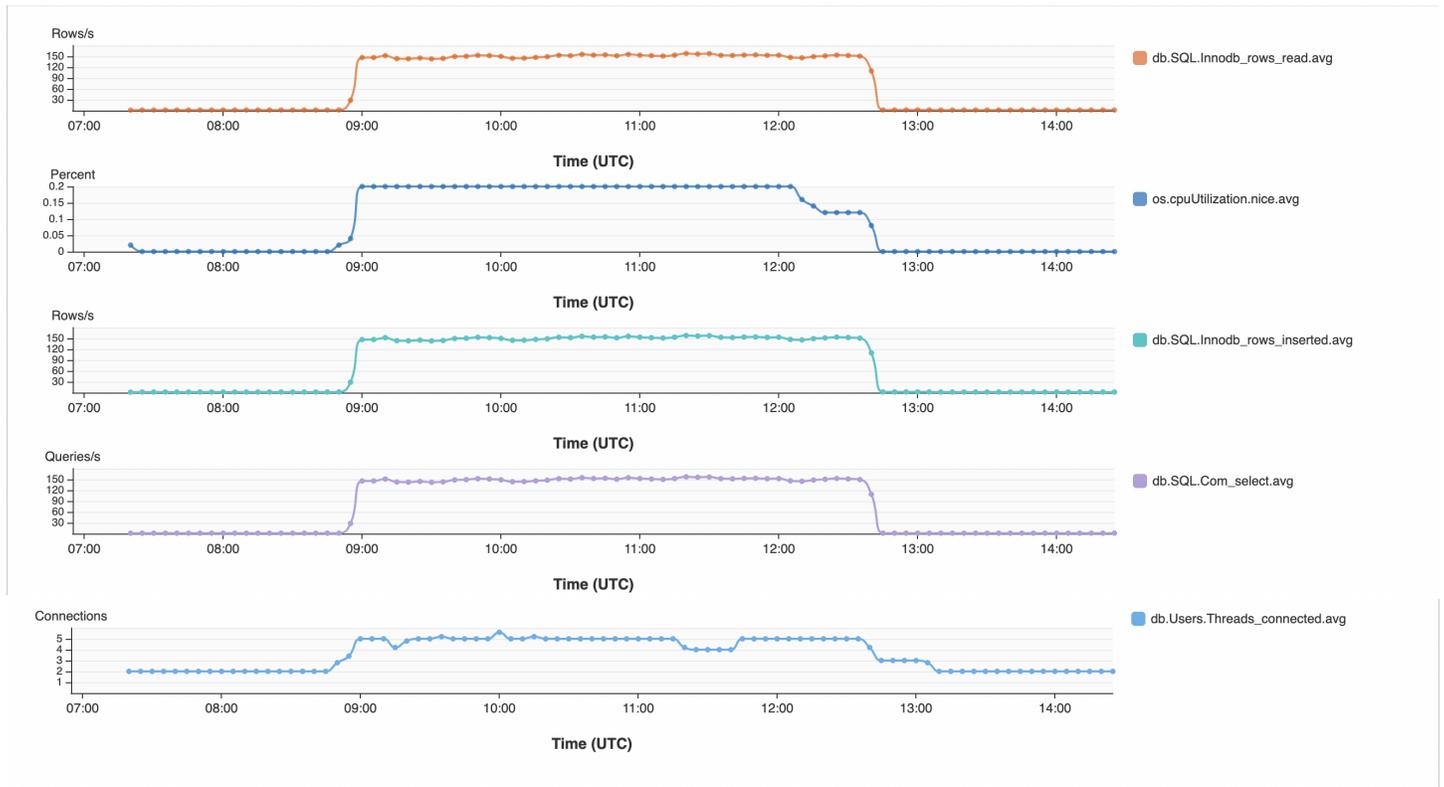
300,024 行をアーカイブするのに約 34 分かかりました。このテーブルには 200 万行ありましたが、ツールはインデックス付きの列に固有のデータを含む行のみをアーカイブしました。

主キーを持つテーブルをアーカイブする (データベースに負荷をかけない)

```
Started at 2022-11-16T08:53:49, ended at 2022-11-16T12:38:18
Action      Count          Time           Pct
commit      4000000       11065.9534    82.16
select      2000000       1278.1854     9.49
inserting   1999999       1050.4961     7.80
other       0              74.1519       0.55
```

1,999,999 行をアーカイブするのに約 3 時間 44 分 29 秒かかりました。

次のグラフは、pt-archiver をシステムに負荷をかけずに単独で実行した場合、CPU とリソースをほとんど消費しないことを示しています。



主キーを持つアーカイブテーブル (データベースに負荷がかかる)

Started at 2022-11-16T17:37:07, ended at 2022-11-17T03:20:43

Action	Count	Time	Pct
commit	4000000	19688.8362	56.23
inserting	1999999	13933.4418	39.79
select	2000000	1305.1770	3.73
other	0	89.1787	0.25

1999999行をアーカイブするのに約9時間、43分、36秒かかりました。

次のグラフは、テスト中に sysbench の負荷がかかったために CPU 使用率が最大 15% に達したことを示しています。ロードが完了した後も、pt-archiver は期待どおり最小限の CPU を消費して処理を続け、アーカイブを完了しました。



グラフから明らかなように、pt-archiverはデータベースに負荷がかかっても積極的にアーカイブしません。

ドキュメント履歴

このガイドは、このドキュメントの大きな変更点をまとめたものです。今後の更新に関する通知を受け取る場合は、[RSS フィード](#)をサブスクライブできます。

変更	説明	日付
初版発行	—	2023 年 6 月 8 日

AWS 規範的ガイドの用語集

以下は、AWS 規範的ガイドが提供する戦略、ガイド、パターンで一般的に使用される用語です。エントリを提案するには、用語集の最後のフィードバックの提供リンクを使用します。

数字

7 Rs

アプリケーションをクラウドに移行するための7つの一般的な移行戦略。これらの戦略は、ガートナーが2011年に特定した5Rsに基づいて構築され、以下で構成されています。

- リファクタリング/アーキテクチャの再設計 — クラウドネイティブ特徴を最大限に活用して、俊敏性、パフォーマンス、スケーラビリティを向上させ、アプリケーションを移動させ、アーキテクチャを変更します。これには、通常、オペレーティングシステムとデータベースの移植が含まれます。例: オンプレミスの Oracle データベースを Amazon Aurora PostgreSQL 互換エディションに移行します。
- リプラットフォーム (リフトアンドリシェイプ) — アプリケーションをクラウドに移行し、クラウド機能を活用するための最適化レベルを導入します。例: オンプレミスの Oracle データベースをの Oracle 用 Amazon Relational Database Service (Amazon RDS) に移行します AWS クラウド。
- 再購入 (ドロップアンドショップ) — 通常、従来のライセンスから SaaS モデルに移行して、別の製品に切り替えます。例: カスタマーリレーションシップ管理 (CRM) システムを Salesforce.com に移行します。
- リホスト (リフトアンドシフト) — クラウド機能を活用するための変更を加えずに、アプリケーションをクラウドに移行します。例: オンプレミスの Oracle データベースをの EC2 インスタンス上の Oracle に移行します AWS クラウド。
- 再配置 (ハイパーバイザーレベルのリフトアンドシフト) — 新しいハードウェアを購入したり、アプリケーションを書き換えたり、既存の運用を変更したりすることなく、インフラストラクチャをクラウドに移行できます。サーバーをオンプレミスプラットフォームから同じプラットフォームのクラウドサービスに移行します。例: Microsoft Hyper-Vアプリケーションをに移行します AWS。
- 保持 (再アクセス) — アプリケーションをお客様のソース環境で保持します。これには、主要なリファクタリングを必要とするアプリケーションや、お客様がその作業を後日まで延期したいアプリケーション、およびそれら移行するためのビジネス上の正当性がないため、お客様が保持するレガシーアプリケーションなどがあります。

- 使用停止 — お客様のソース環境で不要になったアプリケーションを停止または削除します。

A

ABAC

[「属性ベースのアクセスコントロール」](#)を参照してください。

抽象化されたサービス

[「マネージドサービス」](#)を参照してください。

ACID

[「原子性、一貫性、分離性、耐久性」](#)を参照してください。

アクティブ - アクティブ移行

(双方向レプリケーションツールまたは二重書き込み操作を使用して) ソースデータベースとターゲットデータベースを同期させ、移行中に両方のデータベースが接続アプリケーションからのトランザクションを処理するデータベース移行方法。この方法では、1 回限りのカットオーバーの必要がなく、管理された小規模なバッチで移行できます。アクティブ [/パッシブ移行](#)よりも柔軟ですが、より多くの作業が必要です。

アクティブ - パッシブ移行

ソースデータベースとターゲットデータベースを同期させながら、データがターゲットデータベースにレプリケートされている間、接続しているアプリケーションからのトランザクションをソースデータベースのみで処理するデータベース移行の方法。移行中、ターゲットデータベースはトランザクションを受け付けません。

集計関数

行のグループを操作し、グループの単一の戻り値を計算する SQL 関数。集計関数の例としては、SUMや などがありますMAX。

AI

[「人工知能」](#)を参照してください。

AI Ops

[「人工知能オペレーション」](#)を参照してください。

匿名化

データセット内の個人情報を完全に削除するプロセス。匿名化は個人のプライバシー保護に役立ちます。匿名化されたデータは、もはや個人データとは見なされません。

アンチパターン

繰り返し起こる問題に対して頻繁に用いられる解決策で、その解決策が逆効果であったり、効果がなかったり、代替案よりも効果が低かったりするもの。

アプリケーションコントロール

マルウェアからシステムを保護するために、承認されたアプリケーションのみを使用できるようにするセキュリティアプローチ。

アプリケーションポートフォリオ

アプリケーションの構築と維持にかかるコスト、およびそのビジネス価値を含む、組織が使用する各アプリケーションに関する詳細情報の集まり。この情報は、[ポートフォリオの検出と分析プロセス](#)の需要要素であり、移行、モダナイズ、最適化するアプリケーションを特定し、優先順位を付けるのに役立ちます。

人工知能 (AI)

コンピューティングテクノロジーを使用し、学習、問題の解決、パターンの認識など、通常は人間に関連づけられる認知機能の実行に特化したコンピュータサイエンスの分野。詳細については、「[人工知能 \(AI\) とは何ですか?](#)」を参照してください。

AI オペレーション (AIOps)

機械学習技術を使用して運用上の問題を解決し、運用上のインシデントと人の介入を減らし、サービス品質を向上させるプロセス。AWS 移行戦略での AIOps の使用方法については、[オペレーション統合ガイド](#)を参照してください。

非対称暗号化

暗号化用のパブリックキーと復号用のプライベートキーから成る 1 組のキーを使用した、暗号化のアルゴリズム。パブリックキーは復号には使用されないため共有しても問題ありませんが、プライベートキーの利用は厳しく制限する必要があります。

原子性、一貫性、分離性、耐久性 (ACID)

エラー、停電、その他の問題が発生した場合でも、データベースのデータ有効性と運用上の信頼性を保証する一連のソフトウェアプロパティ。

属性ベースのアクセス制御 (ABAC)

部署、役職、チーム名など、ユーザーの属性に基づいてアクセス許可をきめ細かく設定する方法。詳細については、AWS Identity and Access Management (IAM) ドキュメントの「[ABAC AWS](#)」を参照してください。

信頼できるデータソース

最も信頼性のある情報源とされるデータのプライマリーバージョンを保存する場所。匿名化、編集、仮名化など、データを処理または変更する目的で、信頼できるデータソースから他の場所にデータをコピーすることができます。

アベイラビリティゾーン

他のアベイラビリティゾーンの障害から AWS リージョン 隔離され、同じリージョン内の他のアベイラビリティゾーンへの低コストで低レイテンシーのネットワーク接続を提供する 内の別の場所。

AWS クラウド導入フレームワーク (AWS CAF)

組織がクラウドに正常に移行 AWS するための効率的で効果的な計画を立てるのに役立つ、のガイドラインとベストプラクティスのフレームワーク。AWS CAF は、ビジネス、人材、ガバナンス、プラットフォーム、セキュリティ、運用という 6 つの重点分野にガイダンスを編成します。ビジネス、人材、ガバナンスの観点では、ビジネススキルとプロセスに重点を置き、プラットフォーム、セキュリティ、オペレーションの視点は技術的なスキルとプロセスに焦点を当てています。例えば、人材の観点では、人事 (HR)、人材派遣機能、および人材管理を扱うステークホルダーを対象としています。この観点から、AWS CAF は、クラウド導入を成功させるための組織の準備に役立つ人材開発、トレーニング、コミュニケーションに関するガイダンスを提供します。詳細については、[AWS CAF ウェブサイト](#) と [AWS CAF のホワイトペーパー](#) を参照してください。

AWS ワークロード認定フレームワーク (AWS WQF)

データベース移行ワークロードを評価し、移行戦略を推奨し、作業見積もりを提供するツール。AWS WQF は AWS Schema Conversion Tool (AWS SCT) に含まれています。データベーススキーマとコードオブジェクト、アプリケーションコード、依存関係、およびパフォーマンス特性を分析し、評価レポートを提供します。

B

不正なボット

個人または組織に混乱や損害を与えることを目的とした[ボット](#)。

BCP

[事業継続計画を参照してください](#)。

動作グラフ

リソースの動作とインタラクションを経時的に示した、一元的なインタラクティブビュー。Amazon Detective の動作グラフを使用すると、失敗したログオンの試行、不審な API 呼び出し、その他同様のアクションを調べることができます。詳細については、Detective ドキュメントの[Data in a behavior graph](#)を参照してください。

ビッグエンディアンシステム

最上位バイトを最初に格納するシステム。[エンディアンネス](#) も参照してください。

二項分類

バイナリ結果 (2 つの可能なクラスのうちの一つ) を予測するプロセス。例えば、お客様の機械学習モデルで「この E メールはスパムですか、それともスパムではありませんか」などの問題を予測する必要があるかもしれません。または「この製品は書籍ですか、車ですか」などの問題を予測する必要があるかもしれません。

ブルームフィルター

要素がセットのメンバーであるかどうかをテストするために使用される、確率的でメモリ効率の高いデータ構造。

ブルー/グリーンデプロイ

2 つの異なる同一の環境を作成するデプロイ戦略。現在のアプリケーションバージョンは 1 つの環境 (青) で実行し、新しいアプリケーションバージョンは他の環境 (緑) で実行します。この戦略は、影響を最小限に抑えながら迅速にロールバックするのに役立ちます。

ボット

インターネット経由で自動タスクを実行し、人間のアクティビティやインタラクションをシミュレートするソフトウェアアプリケーション。インターネット上の情報のインデックスを作成するウェブクローラーなど、一部のボットは有用または有益です。悪質なボットと呼ばれる他のボット

トの中には、個人や組織に混乱を与えたり、損害を与えたりすることを意図しているものがあります。

ボットネット

[マルウェア](#)に感染し、[ボット](#)のヘルダーまたはボットオペレーターと呼ばれる、単一関係者の管理下にあるボットのネットワーク。ボットは、ボットとその影響をスケールするための最もよく知られているメカニズムです。

ブランチ

コードリポジトリに含まれる領域。リポジトリに最初に作成するブランチは、メインブランチといます。既存のブランチから新しいブランチを作成し、その新しいブランチで機能を開発したり、バグを修正したりできます。機能を構築するために作成するブランチは、通常、機能ブランチと呼ばれます。機能をリリースする準備ができたなら、機能ブランチをメインブランチに統合します。詳細については、「[ブランチについて](#) (GitHub ドキュメント)」を参照してください。

ブレイクグラスアクセス

例外的な状況や承認されたプロセスを通じて、ユーザーが通常アクセス許可を持たない AWS アカウント にすばやくアクセスできるようにします。詳細については、Well-Architected [ガイド](#)の「[ブレイクグラス手順の実装](#)」インジケータを参照してください。AWS

ブラウフィールド戦略

環境の既存インフラストラクチャ。システムアーキテクチャにブラウフィールド戦略を導入する場合、現在のシステムとインフラストラクチャの制約に基づいてアーキテクチャを設計します。既存のインフラストラクチャを拡張している場合は、ブラウフィールド戦略と[グリーンフィールド](#)戦略を融合させることもできます。

バッファキャッシュ

アクセス頻度が最も高いデータが保存されるメモリ領域。

ビジネス能力

価値を生み出すためにビジネスが行うこと (営業、カスタマーサービス、マーケティングなど)。マイクロサービスのアーキテクチャと開発の決定は、ビジネス能力によって推進できます。詳細については、ホワイトペーパー [AWSでのコンテナ化されたマイクロサービスの実行](#) の [ビジネス機能を中心に組織化](#) セクションを参照してください。

ビジネス継続性計画 (BCP)

大規模移行など、中断を伴うイベントが運用に与える潜在的な影響に対処し、ビジネスを迅速に再開できるようにする計画。

C

CAF

[AWS 「クラウド導入フレームワーク」を参照してください。](#)

Canary デプロイ

エンドユーザーへのバージョンの低速かつ増分的なリリース。確信できたら、新しいバージョンをデプロイし、現在のバージョン全体を置き換えます。

CCoE

[「Cloud Center of Excellence」を参照してください。](#)

CDC

[「データキャプチャの変更」を参照してください。](#)

変更データキャプチャ (CDC)

データソース (データベーステーブルなど) の変更を追跡し、その変更に関するメタデータを記録するプロセス。CDC は、ターゲットシステムでの変更を監査またはレプリケートして同期を維持するなど、さまざまな目的に使用できます。

カオスエンジニアリング

障害や破壊的なイベントを意図的に導入して、システムの耐障害性をテストします。[AWS Fault Injection Service \(AWS FIS \)](#) を使用して、AWS ワークロードに負荷をかけてレスポンスを評価する実験を実行できます。

CI/CD

[「継続的インテグレーションと継続的デリバリー」を参照してください。](#)

分類

予測を生成するのに役立つ分類プロセス。分類問題の機械学習モデルは、離散値を予測します。離散値は、常に互いに区別されます。例えば、モデルがイメージ内に車があるかどうかを評価する必要がある場合があります。

クライアント側の暗号化

ターゲットがデータ AWS のサービスを受信する前に、ローカルでデータを暗号化します。

Cloud Center of Excellence (CCoE)

クラウドのベストプラクティスの作成、リソースの移動、移行のタイムラインの確立、大規模変革を通じて組織をリードするなど、組織全体のクラウド導入の取り組みを推進する学際的なチーム。詳細については、AWS クラウド エンタープライズ戦略ブログの[CCoE の投稿](#)を参照してください。

クラウドコンピューティング

リモートデータストレージと IoT デバイス管理に通常使用されるクラウドテクノロジー。クラウドコンピューティングは、一般的に[エッジコンピューティング](#)テクノロジーに接続されています。

クラウド運用モデル

IT 組織において、1 つ以上のクラウド環境を構築、成熟、最適化するために使用される運用モデル。詳細については、[「クラウド運用モデルの構築」](#)を参照してください。

導入のクラウドステージ

組織が移行するときに通常実行する 4 つのフェーズ AWS クラウド :

- プロジェクト — 概念実証と学習を目的として、クラウド関連のプロジェクトをいくつか実行する
- 基礎固め — お客様のクラウドの導入を拡大するための基礎的な投資 (ランディングゾーン の作成、CCoE の定義、運用モデルの確立など)
- 移行 — 個々のアプリケーションの移行
- 再発明 — 製品とサービスの最適化、クラウドでのイノベーション

これらのステージは、AWS クラウド エンタープライズ戦略ブログのブログ記事[「クラウドファーストへのジャーニー」](#)と[「導入のステージ」](#)で Stephen Orban によって定義されました。移行戦略とどのように関連しているかについては、AWS [「移行準備ガイド」](#)を参照してください。

CMDB

[「設定管理データベース」](#)を参照してください。

コードリポジトリ

ソースコードやその他の資産 (ドキュメント、サンプル、スクリプトなど) が保存され、バージョン管理プロセスを通じて更新される場所。一般的なクラウドリポジトリには、GitHub またはが含まれます AWS CodeCommit。コードの各バージョンはブランチと呼ばれます。マイクロサー

ビスの構造では、各リポジトリは 1 つの機能専用です。1 つの CI/CD パイプラインで複数のリポジトリを使用できます。

コールドキャッシュ

空である、または、かなり空きがある、もしくは、古いデータや無関係なデータが含まれているバッファキャッシュ。データベースインスタンスはメインメモリまたはディスクから読み取る必要があり、バッファキャッシュから読み取るよりも時間がかかるため、パフォーマンスに影響します。

コールドデータ

めったにアクセスされず、通常は過去のデータです。この種類のデータをクエリする場合、通常は低速なクエリでも問題ありません。このデータを低パフォーマンスで安価なストレージ階層またはクラスに移動すると、コストを削減することができます。

コンピュータビジョン (CV)

機械学習を使用してデジタルイメージやビデオなどのビジュアル形式から情報を分析および抽出する [AI](#) の分野。例えば、はオンプレミスのカメラネットワークに CV を追加するデバイス AWS Panorama を提供し、Amazon SageMaker は CV の画像処理アルゴリズムを提供します。

設定ドリフト

ワークロードの場合、設定は想定した状態から変化します。これにより、ワークロードが非準拠になる可能性があり、通常は段階的かつ意図的ではありません。

構成管理データベース (CMDB)

データベースとその IT 環境 (ハードウェアとソフトウェアの両方のコンポーネントとその設定を含む) に関する情報を保存、管理するリポジトリ。通常、CMDB のデータは、移行のポートフォリオの検出と分析の段階で使用します。

コンフォーマンスパック

コンプライアンスチェックとセキュリティチェックをカスタマイズするためにアセンブルできる AWS Config ルールと修復アクションのコレクション。YAML テンプレートを使用して、コンフォーマンスパックを AWS アカウント および リージョンで単一のエンティティとしてデプロイすることも、組織全体にデプロイすることもできます。詳細については、AWS Config ドキュメントの「[コンフォーマンスパック](#)」を参照してください。

継続的インテグレーションと継続的デリバリー (CI/CD)

ソフトウェアリリースプロセスのソース、ビルド、テスト、ステージング、本番の各ステージを自動化するプロセス。CI/CD は一般的にパイプラインと呼ばれます。プロセスの自動化、生産性

の向上、コード品質の向上、配信の加速化を可能にします。詳細については、「[継続的デリバリーの利点](#)」を参照してください。CD は継続的デプロイ (Continuous Deployment) の略語でもあります。詳細については「[継続的デリバリーと継続的なデプロイ](#)」を参照してください。

CV

[「コンピュータビジョン」](#)を参照してください。

D

保管中のデータ

ストレージ内にあるデータなど、常に自社のネットワーク内にあるデータ。

データ分類

ネットワーク内のデータを重要度と機密性に基づいて識別、分類するプロセス。データに適した保護および保持のコントロールを判断する際に役立つため、あらゆるサイバーセキュリティのリスク管理戦略において重要な要素です。データ分類は、AWS Well-Architected フレームワークのセキュリティの柱のコンポーネントです。詳細については、[データ分類](#)を参照してください。

データドリフト

実稼働データと ML モデルのトレーニングに使用されたデータとの間に有意な差異が生じたり、入力データが時間の経過と共に有意に変化したりすることです。データドリフトは、ML モデル予測の全体的な品質、精度、公平性を低下させる可能性があります。

転送中のデータ

ネットワーク内 (ネットワークリソース間など) を活発に移動するデータ。

データメッシュ

一元化された管理とガバナンスにより、分散型の分散型データ所有権を提供するアーキテクチャフレームワーク。

データ最小化

厳密に必要なデータのみを収集し、処理するという原則。でデータ最小化を実践 AWS クラウドすることで、プライバシーリスク、コスト、分析のカーボンフットプリントを削減できます。

データ境界

AWS 環境内の一連の予防ガードレール。信頼できる ID のみが、期待されるネットワークから信頼できるリソースにアクセスしていることを確認できます。詳細については、「[「でのデータ境界の構築 AWS」](#)」を参照してください。

データの前処理

raw データをお客様の機械学習モデルで簡単に解析できる形式に変換すること。データの前処理とは、特定の列または行を削除して、欠落している、矛盾している、または重複する値に対処することを意味します。

データ出所

データの生成、送信、保存の方法など、データのライフサイクル全体を通じてデータの出所と履歴を追跡するプロセス。

データ件名

データを収集、処理している個人。

データウェアハウス

分析などのビジネスインテリジェンスをサポートするデータ管理システム。データウェアハウスには通常、大量の履歴データが含まれており、クエリや分析によく使用されます。

データベース定義言語 (DDL)

データベース内のテーブルやオブジェクトの構造を作成または変更するためのステートメントまたはコマンド。

データベース操作言語 (DML)

データベース内の情報を変更 (挿入、更新、削除) するためのステートメントまたはコマンド。

DDL

[「データベース定義言語」](#)を参照してください。

ディープアンサンブル

予測のために複数の深層学習モデルを組み合わせる。ディープアンサンブルを使用して、より正確な予測を取得したり、予測の不確実性を推定したりできます。

ディープラーニング

人工ニューラルネットワークの複数層を使用して、入力データと対象のターゲット変数の間のマッピングを識別する機械学習サブフィールド。

defense-in-depth

一連のセキュリティメカニズムとコントロールをコンピュータネットワーク全体に層状に重ねて、ネットワークとその内部にあるデータの機密性、整合性、可用性を保護する情報セキュリティの手法。この戦略を に採用するときは AWS、AWS Organizations 構造の異なるレイヤーに複数のコントロールを追加して、リソースの安全性を確保します。例えば、defense-in-depth アプローチでは、多要素認証、ネットワークセグメンテーション、暗号化を組み合わせることができます。

委任管理者

では AWS Organizations、互換性のあるサービスが AWS メンバーアカウントを登録して組織のアカウントを管理し、そのサービスのアクセス許可を管理できます。このアカウントを、そのサービスの委任管理者と呼びます。詳細、および互換性のあるサービスの一覧は、AWS Organizations ドキュメントの [AWS Organizationsで利用できるサービス](#) を参照してください。

デプロイメント

アプリケーション、新機能、コードの修正をターゲットの環境で利用できるようにするプロセス。デプロイでは、コードベースに変更を施した後、アプリケーションの環境でそのコードベースを構築して実行します。

開発環境

[「環境」](#) を参照してください。

検出管理

イベントが発生したときに、検出、ログ記録、警告を行うように設計されたセキュリティコントロール。これらのコントロールは副次的な防衛手段であり、実行中の予防的コントロールをすり抜けたセキュリティイベントをユーザーに警告します。詳細については、Implementing security controls on AWSの [Detective controls](#) を参照してください。

開発バリューストリームマッピング (DVSM)

ソフトウェア開発ライフサイクルのスピードと品質に悪影響を及ぼす制約を特定し、優先順位を付けるために使用されるプロセス。DVSM は、もともとリーンマニユファクチャリング・プラクティスのために設計されたバリューストリームマッピング・プロセスを拡張したものです。ソフトウェア開発プロセスを通じて価値を創造し、動かすために必要なステップとチームに焦点を当てています。

デジタルツイン

建物、工場、産業機器、生産ラインなど、現実世界のシステムを仮想的に表現したものです。デジタルツインは、予知保全、リモートモニタリング、生産最適化をサポートします。

ディメンションテーブル

[スタースキーマ](#) では、ファクトテーブル内の量的データに関するデータ属性を含む小さなテーブル。ディメンションテーブル属性は通常、テキストフィールドまたはテキストのように動作する離散数値です。これらの属性は、クエリの制約、フィルタリング、結果セットのラベル付けに一般的に使用されます。

ディザスタ

ワークロードまたはシステムが、導入されている主要な場所でのビジネス目標の達成を妨げるイベント。これらのイベントは、自然災害、技術的障害、または意図しない設定ミスやマルウェア攻撃などの人間の行動の結果である場合があります。

ディザスタリカバリ (DR)

[災害](#)によるダウンタイムとデータ損失を最小限に抑えるために使用する戦略とプロセス。詳細については、AWS Well-Architected [フレームワークの「でのワークロードのディザスタリカバリ」](#) [AWS: クラウドでのリカバリ](#) を参照してください。

DML

[「データベース操作言語」](#) を参照してください。

ドメイン駆動型設計

各コンポーネントが提供している変化を続けるドメイン、またはコアビジネス目標にコンポーネントを接続して、複雑なソフトウェアシステムを開発するアプローチ。この概念は、エリック・エヴァンスの著書、Domain-Driven Design: Tackling Complexity in the Heart of Software (ドメイン駆動設計:ソフトウェアの中心における複雑さへの取り組み) で紹介されています (ボストン: Addison-Wesley Professional、2003)。strangler fig パターンでドメイン駆動型設計を使用する方法の詳細については、[コンテナと Amazon API Gateway を使用して、従来の Microsoft ASP.NET \(ASMX\) ウェブサービスを段階的にモダナイズ](#) を参照してください。

DR

[「ディザスタリカバリ」](#) を参照してください。

ドリフト検出

ベースライン設定からの偏差の追跡。例えば、AWS CloudFormation を使用して[システムリソースのドリフトを検出したり](#)、を使用して AWS Control Tower ガバナンス要件への準拠に影響を与える可能性のある[ランディングゾーンの変更を検出したり](#)できます。

DVSM

[「開発値ストリームマッピング」](#)を参照してください。

E

EDA

[「探索的データ分析」](#)を参照してください。

エッジコンピューティング

IoT ネットワークのエッジにあるスマートデバイスの計算能力を高めるテクノロジー。[クラウドコンピューティング](#)と比較すると、エッジコンピューティングは通信レイテンシーを短縮し、応答時間を短縮できます。

暗号化

人間が読み取り可能なプレーンテキストデータを暗号文に変換するコンピューティングプロセス。

暗号化キー

暗号化アルゴリズムが生成した、ランダム化されたビットからなる暗号文字列。キーの長さは決まっておらず、各キーは予測できないように、一意になるように設計されています。

エンディアン

コンピュータメモリにバイトが格納される順序。ビッグエンディアンシステムでは、最上位バイトが最初に格納されます。リトルエンディアンシステムでは、最下位バイトが最初に格納されます。

エンドポイント

[「サービスエンドポイント」](#)を参照してください。

エンドポイントサービス

仮想プライベートクラウド (VPC) 内でホストして、他のユーザーと共有できるサービス。を使用してエンドポイントサービスを作成し AWS PrivateLink、他の AWS アカウント または AWS Identity and Access Management (IAM) プリンシパルにアクセス許可を付与できます。これらのアカウントまたはプリンシパルは、インターフェイス VPC エンドポイントを作成することで、エンドポイントサービスにプライベートに接続できます。詳細については、Amazon Virtual Private Cloud (Amazon VPC) ドキュメントの「[エンドポイントサービスを作成する](#)」を参照してください。

エンタープライズリソースプランニング (ERP)

エンタープライズの主要なビジネスプロセス (アカウンティング、[MES](#)、プロジェクト管理など) を自動化および管理するシステム。

エンベロープ暗号化

暗号化キーを、別の暗号化キーを使用して暗号化するプロセス。詳細については、AWS Key Management Service (AWS KMS) [ドキュメントの「エンベロープ暗号化」](#)を参照してください。

環境

実行中のアプリケーションのインスタンス。クラウドコンピューティングにおける一般的な環境の種類は以下のとおりです。

- 開発環境 — アプリケーションのメンテナンスを担当するコアチームのみが利用できる、実行中のアプリケーションのインスタンス。開発環境は、上位の環境に昇格させる変更をテストするときに使用します。このタイプの環境は、テスト環境と呼ばれることもあります。
- 下位環境 — 初期ビルドやテストに使用される環境など、アプリケーションのすべての開発環境。
- 本番環境 — エンドユーザーがアクセスできる、実行中のアプリケーションのインスタンス。CI/CD パイプラインでは、本番環境が最後のデプロイ環境になります。
- 上位環境 — コア開発チーム以外のユーザーがアクセスできるすべての環境。これには、本番環境、本番前環境、ユーザー承認テスト環境などが含まれます。

エピック

アジャイル方法論で、お客様の作業の整理と優先順位付けに役立つ機能カテゴリ。エピックでは、要件と実装タスクの概要についてハイレベルな説明を提供します。例えば、AWS CAF セキュリティエピックには、ID とアクセスの管理、検出コントロール、インフラストラクチャセ

セキュリティ、データ保護、インシデント対応が含まれます。AWS 移行戦略のエピックの詳細については、[プログラム実装ガイド](#) を参照してください。

ERP

[「エンタープライズリソース計画」](#) を参照してください。

探索的データ分析 (EDA)

データセットを分析してその主な特性を理解するプロセス。お客様は、データを収集または集計してから、パターンの検出、異常の検出、および前提条件のチェックのための初期調査を実行します。EDA は、統計の概要を計算し、データの可視化を作成することによって実行されます。

F

ファクトテーブル

[スタースキーマ](#) の中央テーブル。事業運営に関する定量的データを保存します。通常、ファクトテーブルには、メジャーを含む列とディメンションテーブルへの外部キーを含む列の 2 種類の列が含まれます。

フェイルファスト

頻繁で段階的なテストを使用して開発ライフサイクルを短縮する哲学。これはアジャイルアプローチの重要な部分です。

障害分離境界

では AWS クラウド、障害の影響を制限し、ワークロードの耐障害性を向上させるアベイラビリティゾーン AWS リージョン、コントロールプレーン、データプレーンなどの境界です。詳細については、[AWS 「障害分離境界」](#) を参照してください。

機能ブランチ

[「ブランチ」](#) を参照してください。

特徴量

お客様が予測に使用する入力データ。例えば、製造コンテキストでは、特徴量は製造ラインから定期的にキャプチャされるイメージの可能性もあります。

特徴量重要度

モデルの予測に対する特徴量の重要性。これは通常、Shapley Additive Deskonations (SHAP) や積分勾配など、さまざまな手法で計算できる数値スコアで表されます。詳細については、[「を使用した機械学習モデルの解釈可能性：AWS」](#)を参照してください。

機能変換

追加のソースによるデータのエンリッチ化、値のスケーリング、単一のデータフィールドからの複数の情報セットの抽出など、機械学習プロセスのデータを最適化すること。これにより、機械学習モデルはデータの恩恵を受けることができます。例えば、「2021-05-27 00:15:37」の日付を「2021年」、「5月」、「木」、「15」に分解すると、学習アルゴリズムがさまざまなデータコンポーネントに関連する微妙に異なるパターンを学習するのに役立ちます。

FGAC

[「きめ細かなアクセスコントロール」](#)を参照してください。

きめ細かなアクセス制御 (FGAC)

複数の条件を使用してアクセス要求を許可または拒否すること。

フラッシュカット移行

段階的なアプローチを使用するのではなく、[変更データキャプチャ](#)による継続的なデータレプリケーションを使用して、可能な限り短い時間でデータを移行するデータベース移行方法。目的はダウンタイムを最小限に抑えることです。

G

ジオブロッキング

[「地理的制限」](#)を参照してください。

地理的制限 (ジオブロッキング)

Amazon では CloudFront、特定の国のユーザーがコンテンツディストリビューションにアクセスできないようにするオプションです。アクセスを許可する国と禁止する国は、許可リストまたは禁止リストを使って指定します。詳細については、CloudFront ドキュメントの[「コンテンツの地理的ディストリビューションの制限」](#)を参照してください。

Gitflow ワークフロー

下位環境と上位環境が、ソースコードリポジトリでそれぞれ異なるブランチを使用する方法。Gitflow ワークフローはレガシーと見なされ、[トランクベースのワークフロー](#)はモダンで推奨されるアプローチです。

グリーンフィールド戦略

新しい環境に既存のインフラストラクチャが存在しないこと。システムアーキテクチャにグリーンフィールド戦略を導入する場合、既存のインフラストラクチャ (別名 [ブラウンフィールド](#)) との互換性の制約を受けることなく、あらゆる新しいテクノロジーを選択できます。既存のインフラストラクチャを拡張している場合は、ブラウンフィールド戦略とグリーンフィールド戦略を融合させることもできます。

ガードレール

組織単位 (OU) 全般のリソース、ポリシー、コンプライアンスを管理するのに役立つ概略的なルール。予防ガードレールは、コンプライアンス基準に一致するようにポリシーを実施します。これらは、サービスコントロールポリシーと IAM アクセス許可の境界を使用して実装されます。検出ガードレールは、ポリシー違反やコンプライアンス上の問題を検出し、修復のためのアラートを発信します。これらは、AWS Config、Amazon AWS Security Hub、GuardDuty、Amazon Inspector AWS Trusted Advisor、およびカスタム AWS Lambda チェックを使用して実装されます。

H

HA

[「高可用性」](#)を参照してください。

異種混在データベースの移行

別のデータベースエンジンを使用するターゲットデータベースへお客様の出典データベースの移行 (例えば、Oracle から Amazon Aurora)。異種間移行は通常、アーキテクチャの再設計作業の一部であり、スキーマの変換は複雑なタスクになる可能性があります。[AWS は、スキーマの変換に役立つ AWS SCTを提供します。](#)

ハイアベイラビリティ (HA)

課題や災害が発生した場合に、介入なしにワークロードを継続的に運用できること。HA システムは、自動的にフェイルオーバーし、一貫して高品質のパフォーマンスを提供し、パフォーマンスへの影響を最小限に抑えながらさまざまな負荷や障害を処理するように設計されています。

ヒストリアンのモダナイゼーション

製造業のニーズによりよく応えるために、オペレーションテクノロジー (OT) システムをモダナイズし、アップグレードするためのアプローチ。ヒストリアンは、工場内のさまざまなソースからデータを収集して保存するために使用されるデータベースの一種です。

同種データベースの移行

お客様の出典データベースを、同じデータベースエンジンを共有するターゲットデータベース (Microsoft SQL Server から Amazon RDS for SQL Server など) に移行する。同種間移行は、通常、リホストまたはリプラットフォーム化の作業の一部です。ネイティブデータベースユーティリティを使用して、スキーマを移行できます。

ホットデータ

リアルタイムデータや最近の翻訳データなど、頻繁にアクセスされるデータ。通常、このデータには高速なクエリ応答を提供する高性能なストレージ階層またはクラスが必要です。

ホットフィックス

本番環境の重大な問題を修正するために緊急で配布されるプログラム。緊急性のため、通常、修正は一般的な DevOps リリースワークフローの外で行われます。

ハイパーケア期間

カットオーバー直後、移行したアプリケーションを移行チームがクラウドで管理、監視して問題に対処する期間。通常、この期間は 1~4 日です。ハイパーケア期間が終了すると、アプリケーションに対する責任は一般的に移行チームからクラウドオペレーションチームに移ります。

IaC

[「Infrastructure as Code」](#) を参照してください。

ID ベースのポリシー

AWS クラウド 環境内のアクセス許可を定義する 1 つ以上の IAM プリンシパルにアタッチされたポリシー。

アイドル状態のアプリケーション

90 日間の平均的な CPU およびメモリ使用率が 5~20% のアプリケーション。移行プロジェクトでは、これらのアプリケーションを廃止するか、オンプレミスに保持するのが一般的です。

IIoT

[「産業モノのインターネット」を参照してください。](#)

イミュータブルインフラストラクチャ

既存のインフラストラクチャを更新、パッチ適用、または変更するのではなく、本番ワークロード用の新しいインフラストラクチャをデプロイするモデル。イミュータブルインフラストラクチャは、[本質的にミュータブルインフラストラクチャ](#) よりも一貫性、信頼性、予測性が高くなります。詳細については、AWS Well-Architected フレームワークの[「イミュータブルインフラストラクチャを使用したデプロイ」](#)のベストプラクティスを参照してください。

インバウンド (受信) VPC

AWS マルチアカウントアーキテクチャでは、アプリケーションの外部からネットワーク接続を受け入れ、検査し、ルーティングする VPC。[AWS Security Reference Architecture](#) では、アプリケーションとより広範なインターネット間の双方向のインターフェイスを保護するために、インバウンド、アウトバウンド、インスペクションの各 VPC を使用してネットワークアカウントを設定することを推奨しています。

増分移行

アプリケーションを 1 回ですべてカットオーバーするのではなく、小さい要素に分けて移行するカットオーバー戦略。例えば、最初は少数のマイクロサービスまたはユーザーのみを新しいシステムに移行する場合があります。すべてが正常に機能することを確認できたら、残りのマイクロサービスやユーザーを段階的に移行し、レガシーシステムを廃止できるようにします。この戦略により、大規模な移行に伴うリスクが軽減されます。

インダストリー 4.0

接続、リアルタイムデータ、自動化、分析、AI/ML の進歩を通じて、のビジネスプロセスのモダナイズを指すために 2016 年に [Klaus Schwab](#) によって導入された用語。

インフラストラクチャ

アプリケーションの環境に含まれるすべてのリソースとアセット。

Infrastructure as Code (IaC)

アプリケーションのインフラストラクチャを一連の設定ファイルを使用してプロビジョニングし、管理するプロセス。IaC は、新しい環境を再現可能で信頼性が高く、一貫性のあるものにするため、インフラストラクチャを一元的に管理し、リソースを標準化し、スケールを迅速に行えるように設計されています。

産業分野における IoT (IIoT)

製造、エネルギー、自動車、ヘルスケア、ライフサイエンス、農業などの産業部門におけるインターネットに接続されたセンサーやデバイスの使用。詳細については、「[Building an industrial Internet of Things \(IIoT\) digital transformation strategy](#)」を参照してください。

インスペクション VPC

AWS マルチアカウントアーキテクチャでは、VPC (同一または異なる 内 AWS リージョン)、インターネット、オンプレミスネットワーク間のネットワークトラフィックの検査を管理する一元化された VPCs。 [AWS Security Reference Architecture](#) では、アプリケーションとより広範なインターネット間の双方向のインターフェイスを保護するために、インバウンド、アウトバウンド、インスペクションの各 VPC を使用してネットワークアカウントを設定することを推奨しています。

IoT

インターネットまたはローカル通信ネットワークを介して他のデバイスやシステムと通信する、センサーまたはプロセッサが組み込まれた接続済み物理オブジェクトのネットワーク。詳細については、「[IoT とは](#)」を参照してください。

解釈可能性

機械学習モデルの特性で、モデルの予測がその入力にどのように依存するかを人間が理解できる度合いを表します。詳細については、「[AWS を使用した機械学習モデルの解釈](#)」を参照してください。

IoT

「[モノのインターネット](#)」を参照してください。

IT 情報ライブラリ (ITIL)

IT サービスを提供し、これらのサービスをビジネス要件に合わせるための一連のベストプラクティス。ITIL は ITSM の基盤を提供します。

IT サービス管理 (ITSM)

組織の IT サービスの設計、実装、管理、およびサポートに関連する活動。クラウドオペレーションと ITSM ツールの統合については、[オペレーション統合ガイド](#) を参照してください。

ITIL

「[IT 情報ライブラリ](#)」を参照してください。

ITSM

[「IT サービス管理」](#)を参照してください。

L

ラベルベースアクセス制御 (LBAC)

強制アクセス制御 (MAC) の実装で、ユーザーとデータ自体にそれぞれセキュリティラベル値が明示的に割り当てられます。ユーザーセキュリティラベルとデータセキュリティラベルが交差する部分によって、ユーザーに表示される行と列が決まります。

ランディングゾーン

ランディングゾーンは、スケーラブルで安全な、適切に設計されたマルチアカウント AWS 環境です。これは、組織がセキュリティおよびインフラストラクチャ環境に自信を持ってワークロードとアプリケーションを迅速に起動してデプロイできる出発点です。ランディングゾーンの詳細については、[安全でスケーラブルなマルチアカウント AWS 環境のセットアップ](#)を参照してください。

大規模な移行

300 台以上のサーバの移行。

LBAC

[「ラベルベースのアクセスコントロール」](#)を参照してください。

最小特権

タスクの実行には必要最低限の権限を付与するという、セキュリティのベストプラクティス。詳細については、IAM ドキュメントの[最小特権アクセス許可を適用する](#)を参照してください。

リフトアンドシフト

[「7R」](#)を参照してください。

リトルエンディアンシステム

最下位バイトを最初に格納するシステム。[エンディアンネス](#)も参照してください。

下位環境

[「環境」](#)を参照してください。

M

機械学習 (ML)

パターン認識と学習にアルゴリズムと手法を使用する人工知能の一種。ML は、モノのインターネット (IoT) データなどの記録されたデータを分析して学習し、パターンに基づく統計モデルを生成します。詳細については、「[機械学習](#)」を参照してください。

メインブランチ

「[ブランチ](#)」を参照してください。

マルウェア

コンピュータのセキュリティまたはプライバシーを侵害するように設計されているソフトウェア。マルウェアは、コンピュータシステムの中断、機密情報の漏洩、不正アクセスにつながる可能性があります。マルウェアの例としては、ウイルス、ワーム、ランサムウェア、トロイの木馬、スパイウェア、キーロガーなどがあります。

マネージドサービス

AWS のサービスがインフラストラクチャレイヤー、オペレーティングシステム、プラットフォーム AWS を運用し、ユーザーがエンドポイントにアクセスしてデータを保存および取得します。Amazon Simple Storage Service (Amazon S3) と Amazon DynamoDB は、マネージドサービスの例です。これらは抽象化されたサービスとも呼ばれます。

製造実行システム (MES)

生産プロセスを追跡、モニタリング、文書化、制御するためのソフトウェアシステム。これにより、加工品を現場の完成製品に変換します。

MAP

「[移行促進プログラム](#)」を参照してください。

メカニズム

ツールを作成し、ツールの導入を推進し、調整のために結果を検査する完全なプロセス。メカニズムとは、動作中にそれ自体を強化して改善するサイクルです。詳細については、AWS Well-Architected フレームワークの「[メカニズムの構築](#)」を参照してください。

メンバーアカウント

内の組織の一部である管理アカウント AWS アカウント を除くすべての AWS Organizations。アカウントが組織のメンバーになることができるのは、一度に 1 つのみです。

MES

[「製造実行システム」](#)を参照してください。

メッセージキューイングテレメトリトランスポート (MQTT)

リソースに制約のある IoT デバイス用の、[パブリッシュ/サブスクライブ](#)パターンに基づく軽量の machine-to-machine (M2M) 通信プロトコル。

マイクロサービス

明確に定義された API を介して通信し、通常は小規模な自己完結型のチームが所有する、小規模で独立したサービスです。例えば、保険システムには、販売やマーケティングなどのビジネス機能、または購買、請求、分析などのサブドメインにマッピングするマイクロサービスが含まれる場合があります。マイクロサービスの利点には、俊敏性、柔軟なスケーリング、容易なデプロイ、再利用可能なコード、回復力などがあります。詳細については、[AWS 「サーバーレスサービスを使用したマイクロサービスの統合」](#)を参照してください。

マイクロサービスアーキテクチャ

各アプリケーションプロセスをマイクロサービスとして実行する独立したコンポーネントを使用してアプリケーションを構築するアプローチ。これらのマイクロサービスは、軽量 API を使用して、明確に定義されたインターフェイスを介して通信します。このアーキテクチャの各マイクロサービスは、アプリケーションの特定の機能に対する需要を満たすように更新、デプロイ、およびスケーリングできます。詳細については、「[でのマイクロサービスの実装 AWS](#)」を参照してください。

Migration Acceleration Program (MAP)

組織がクラウドへの移行のための強固な運用基盤を構築し、移行の初期コストを相殺するのに役立つコンサルティングサポート、トレーニング、サービスを提供する AWS プログラム。MAP には、組織的な方法でレガシー移行を実行するための移行方法論と、一般的な移行シナリオを自動化および高速化する一連のツールが含まれています。

大規模な移行

アプリケーションポートフォリオの大部分を次々にクラウドに移行し、各ウェーブでより多くのアプリケーションを高速に移動させるプロセス。この段階では、以前の段階から学んだベストプラクティスと教訓を使用して、移行ファクトリー チーム、ツール、プロセスのうち、オートメーションとアジャイルデリバリーによってワークロードの移行を合理化します。これは、[AWS 移行戦略](#) の第 3 段階です。

移行ファクトリー

自動化された俊敏性のあるアプローチにより、ワークロードの移行を合理化する部門横断的なチーム。移行ファクトリーチームには、通常、オペレーション、ビジネスアナリストと所有者、移行エンジニア、デベロッパー、スプリントに取り組む DevOps プロフェッショナルが含まれます。エンタープライズアプリケーションポートフォリオの 20~50% は、ファクトリーのアプローチによって最適化できる反復パターンで構成されています。詳細については、このコンテンツセットの[移行ファクトリーに関する解説](#)と[Cloud Migration Factory ガイド](#)を参照してください。

移行メタデータ

移行を完了するために必要なアプリケーションおよびサーバーに関する情報。移行パターンごとに、異なる一連の移行メタデータが必要です。移行メタデータの例には、ターゲットサブネット、セキュリティグループ、AWS アカウントなどがあります。

移行パターン

移行戦略、移行先、および使用する移行アプリケーションまたはサービスを詳述する、反復可能な移行タスク。例: Application Migration Service を使用して Amazon EC2 AWS への移行をリホストします。

Migration Portfolio Assessment (MPA)

に移行するためのビジネスケースを検証するための情報を提供するオンラインツール AWS クラウド。MPA は、詳細なポートフォリオ評価 (サーバーの適切なサイジング、価格設定、TCO 比較、移行コスト分析) および移行プラン (アプリケーションデータの分析とデータ収集、アプリケーションのグループ化、移行の優先順位付け、およびウェーブプランニング) を提供します。[MPA ツール](#) (ログインが必要) は、すべての AWS コンサルタントと APN パートナーコンサルタントが無料で利用できます。

移行準備状況評価 (MRA)

AWS CAF を使用して、組織のクラウド準備状況に関するインサイトを取得し、長所と短所を特定し、特定されたギャップを埋めるためのアクションプランを構築するプロセス。詳細については、[移行準備状況ガイド](#)を参照してください。MRA は、[AWS 移行戦略](#)の第一段階です。

移行戦略

ワークロードを に移行するために使用されるアプローチ AWS クラウド。詳細については、この用語集の「[7 Rs エントリ](#)」と「[組織を動員して大規模な移行を加速する](#)」を参照してください。

ML

[「機械学習」を参照してください。](#)

モダナイゼーション

古い(レガシーまたはモノリシック)アプリケーションとそのインフラストラクチャをクラウド内の俊敏で弾力性のある高可用性システムに変換して、コストを削減し、効率を高め、イノベーションを活用します。詳細については、[「」の「アプリケーションをモダナイズするための戦略 AWS クラウド」](#)を参照してください。

モダナイゼーション準備状況評価

組織のアプリケーションのモダナイゼーションの準備状況を判断し、利点、リスク、依存関係を特定し、組織がこれらのアプリケーションの将来の状態をどの程度適切にサポートできるかを決定するのに役立つ評価。評価の結果として、ターゲットアーキテクチャのブループリント、モダナイゼーションプロセスの開発段階とマイルストーンを詳述したロードマップ、特定されたギャップに対処するためのアクションプランが得られます。詳細については、[「」の「アプリケーションのモダナイゼーション準備状況の評価 AWS クラウド」](#)を参照してください。

モノリシックアプリケーション(モノリス)

緊密に結合されたプロセスを持つ単一のサービスとして実行されるアプリケーション。モノリシックアプリケーションにはいくつかの欠点があります。1つのアプリケーション機能エクスペリエンスの需要が急増する場合は、アーキテクチャ全体をスケーリングする必要があります。モノリシックアプリケーションの特徴を追加または改善することは、コードベースが大きくなると複雑になります。これらの問題に対処するには、マイクロサービスアーキテクチャを使用できます。詳細については、[モノリスをマイクロサービスに分解する](#)を参照してください。

MPA

[「移行ポートフォリオ評価」を参照してください。](#)

MQTT

[「Message Queuing Telemetry Transport」を参照してください。](#)

多クラス分類

複数のクラスの予測を生成するプロセス(2つ以上の結果の1つを予測します)。例えば、機械学習モデルが、「この製品は書籍、自動車、電話のいずれですか?」または、「このお客様にとって最も関心のある商品のカテゴリはどれですか?」と聞くかもしれません。

変更可能なインフラストラクチャ

本番ワークロードの既存のインフラストラクチャを更新および変更するモデル。Well-Architected AWS Framework では、一貫性、信頼性、予測可能性を向上させるために、[イミュータブルなインフラストラクチャ](#)の使用をベストプラクティスとして推奨しています。

O

OAC

[「オリジンアクセスコントロール」](#)を参照してください。

OAI

[「オリジンアクセスアイデンティティ」](#)を参照してください。

OCM

[「組織変更管理」](#)を参照してください。

オフライン移行

移行プロセス中にソースワークロードを停止させる移行方法。この方法はダウンタイムが長くなるため、通常は重要ではない小規模なワークロードに使用されます。

OI

[「オペレーション統合」](#)を参照してください。

OLA

[「運用レベルの契約」](#)を参照してください。

オンライン移行

ソースワークロードをオフラインにせずにターゲットシステムにコピーする移行方法。ワークロードに接続されているアプリケーションは、移行中も動作し続けることができます。この方法はダウンタイムがゼロから最小限で済むため、通常は重要な本番稼働環境のワークロードに使用されます。

OPC-UA

[「Open Process Communications - Unified Architecture」](#)を参照してください。

オープンプロセス通信 - 統合アーキテクチャ (OPC-UA)

産業用オートメーション用の machine-to-machine (M2M) 通信プロトコル。OPC-UA は、データの暗号化、認証、認可スキームを備えた相互運用性標準を提供します。

オペレーショナルレベルアグリーメント (OLA)

サービスレベルアグリーメント (SLA) をサポートするために、どの機能的 IT グループが互いに提供することを約束するかを明確にする契約。

運用準備状況レビュー (ORR)

インシデントや潜在的な障害の理解、評価、防止、または範囲の縮小に役立つ質問とそれに関連するベストプラクティスのチェックリスト。詳細については、AWS Well-Architected フレームワークの「[運用準備状況レビュー \(ORR\)](#)」を参照してください。

運用テクノロジー (OT)

産業運用、機器、インフラストラクチャを制御するために物理環境と連携するハードウェアおよびソフトウェアシステム。製造では、OT と情報技術 (IT) システムの統合が、[Industry 4.0](#) トランスフォーメーションの主要な焦点です。

オペレーション統合 (OI)

クラウドでオペレーションをモダナイズするプロセスには、準備計画、オートメーション、統合が含まれます。詳細については、[オペレーション統合ガイド](#) を参照してください。

組織の証跡

の組織 AWS アカウント 内のすべての のすべてのイベントをログ AWS CloudTrail に記録する、によって作成された証跡 AWS Organizations。証跡は、組織に含まれている各 AWS アカウントに作成され、各アカウントのアクティビティを追跡します。詳細については、ドキュメントの「[組織の証跡の作成](#)」を参照してください。CloudTrail

組織変更管理 (OCM)

人材、文化、リーダーシップの観点から、主要な破壊的なビジネス変革を管理するためのフレームワーク。OCM は、変化の導入を加速し、移行問題に対処し、文化や組織の変化を推進することで、組織が新しいシステムと戦略の準備と移行するのを支援します。AWS 移行戦略では、クラウド導入プロジェクトに必要な変化のスピードから、このフレームワークは人材アクセラレーションと呼ばれます。詳細については、[OCM ガイド](#) を参照してください。

オリジンアクセスコントロール (OAC)

では CloudFront、Amazon Simple Storage Service (Amazon S3) コンテンツを保護するためのアクセスを制限するための拡張オプションです。OAC は、すべての のすべての S3 バケット AWS リージョン、AWS KMS (SSE-KMS) によるサーバー側の暗号化、S3 バケットへの動的 PUT および DELETE リクエストをサポートします。

オリジンアクセスアイデンティティ (OAI)

では CloudFront、Amazon S3 コンテンツを保護するためのアクセスを制限するオプションです。OAI を使用すると、は Amazon S3 が認証できるプリンシパル CloudFront を作成します。認証されたプリンシパルは、特定の CloudFront ディストリビューションを介してのみ S3 バケット内のコンテンツにアクセスできます。[OAC](#)も併せて参照してください。OAC では、より詳細な、強化されたアクセスコントロールが可能です。

ORR

[「運用準備状況レビュー」](#)を参照してください。

OT

[「運用技術」](#)を参照してください。

アウトバウンド (送信) VPC

AWS マルチアカウントアーキテクチャでは、アプリケーション内から開始されるネットワーク接続を処理する VPC。[AWS Security Reference Architecture](#) では、アプリケーションとより広範なインターネット間の双方向のインターフェイスを保護するために、インバウンド、アウトバウンド、インスペクションの各 VPC を使用してネットワークアカウントを設定することを推奨しています。

P

アクセス許可の境界

ユーザーまたはロールが使用できるアクセス許可の上限を設定する、IAM プリンシパルにアタッチされる IAM 管理ポリシー。詳細については、IAM ドキュメントの[アクセス許可の境界](#)を参照してください。

個人を特定できる情報 (PII)

直接閲覧した場合、または他の関連データと組み合わせた場合に、個人の身元を合理的に推測するために使用できる情報。PII の例には、氏名、住所、連絡先情報などがあります。

PII

[「個人を特定できる情報」](#)を参照してください。

プレイブック

クラウドでのコアオペレーション機能の提供など、移行に関連する作業を取り込む、事前定義された一連のステップ。プレイブックは、スクリプト、自動ランブック、またはお客様のモダナイズされた環境を運用するために必要なプロセスや手順の要約などの形式をとることができます。

PLC

[「プログラム可能なロジックコントローラー」](#)を参照してください。

PLM

[「製品ライフサイクル管理」](#)を参照してください。

ポリシー

アクセス許可の定義 ([アイデンティティベースのポリシーを参照](#))、アクセス条件の指定 ([リソースベースのポリシーを参照](#))、またはの組織内のすべてのアカウントに対する最大アクセス許可の定義 AWS Organizations ([サービスコントロールポリシーを参照](#)) が可能なオブジェクト。

多言語の永続性

データアクセスパターンやその他の要件に基づいて、マイクロサービスのデータストレージテクノロジーを個別に選択します。マイクロサービスが同じデータストレージテクノロジーを使用している場合、実装上の問題が発生したり、パフォーマンスが低下する可能性があります。マイクロサービスは、要件に最も適合したデータストアを使用すると、より簡単に実装でき、パフォーマンスとスケーラビリティが向上します。詳細については、[マイクロサービスでのデータ永続性の有効化](#)を参照してください。

ポートフォリオ評価

移行を計画するために、アプリケーションポートフォリオの検出、分析、優先順位付けを行うプロセス。詳細については、「[移行準備状況ガイド](#)」を参照してください。

述語

true または を返すクエリ条件。false 通常は WHERE 句にあります。

述語のプッシュダウン

転送前にクエリ内のデータをフィルタリングするデータベースクエリ最適化手法。これにより、リレーショナルデータベースから取得して処理する必要があるデータの量が減少し、クエリのパフォーマンスが向上します。

予防的コントロール

イベントの発生を防ぐように設計されたセキュリティコントロール。このコントロールは、ネットワークへの不正アクセスや好ましくない変更を防ぐ最前線の防御です。詳細については、Implementing security controls on AWSの[Preventative controls](#)を参照してください。

プリンシパル

アクションを実行し AWS、リソースにアクセスできる のエンティティ。このエンティティは通常、IAM ロール AWS アカウント、またはユーザーのルートユーザーです。詳細については、IAM ドキュメントの[ロールに関する用語と概念](#)内にあるプリンシパルを参照してください。

プライバシーバイデザイン

エンジニアリングプロセス全体を通してプライバシーを考慮に入れたシステムエンジニアリングのアプローチ。

プライベートホストゾーン

1 つ以上の VPC 内のドメインとそのサブドメインへの DNS クエリに対し、Amazon Route 53 がどのように応答するかに関する情報を保持するコンテナ。詳細については、Route 53 ドキュメントの「[プライベートホストゾーンの使用](#)」を参照してください。

プロアクティブコントロール

非準拠のリソースのデプロイを防止するように設計された[セキュリティコントロール](#)。これらのコントロールは、プロビジョニング前にリソースをスキャンします。リソースがコントロールに準拠していない場合、プロビジョニングされません。詳細については、AWS Control Tower ドキュメントの「[コントロールリファレンスガイド](#)」および「[でのセキュリティコントロールの実装](#)」の「[プロアクティブコントロール](#)」を参照してください。 AWS

製品ライフサイクル管理 (PLM)

設計、開発、発売から成長と成熟まで、製品のデータとプロセスのライフサイクル全体にわたる管理、および辞退と削除。

本番環境

[「環境」](#)を参照してください。

プログラミング可能ロジックコントローラー (NAL)

製造では、マシンをモニタリングし、承認プロセスを自動化する、信頼性が高く、適応性の高いコンピュータです。

仮名化

データセット内の個人識別子をプレースホルダー値に置き換えるプロセス。仮名化は個人のプライバシー保護に役立ちます。仮名化されたデータは、依然として個人データとみなされます。

パブリッシュ/サブスクライブ (pub/sub)

マイクロサービス間の非同期通信を可能にするパターン。スケーラビリティと応答性を向上させます。例えば、マイクロサービスベースの [MES](#) では、マイクロサービスは他のマイクロサービスがサブスクライブできるチャンネルにイベントメッセージを発行できます。システムは、公開サービスを変更せずに新しいマイクロサービスを追加できます。

Q

クエリプラン

SQL リレーショナルデータベースシステムのデータにアクセスするために使用される手順などの一連のステップ。

クエリプランのリグレッション

データベースサービスのオプティマイザーが、データベース環境に特定の変更が加えられる前に選択されたプランよりも最適性の低いプランを選択すること。これは、統計、制限事項、環境設定、クエリパラメータのバインディングの変更、およびデータベースエンジンの更新などが原因である可能性があります。

R

RACI マトリックス

[責任、説明責任、相談、通知 \(RACI\)](#) を参照してください。

ランサムウェア

決済が完了するまでコンピュータシステムまたはデータへのアクセスをブロックするように設計された、悪意のあるソフトウェア。

RASCI マトリックス

[責任、説明責任、相談、通知 \(RACI\)](#) を参照してください。

RCAC

[「行と列のアクセスコントロール」を参照してください。](#)

リードレプリカ

読み取り専用で使用されるデータベースのコピー。クエリをリードレプリカにルーティングして、プライマリデータベースへの負荷を軽減できます。

再構築

[「7 Rs」を参照してください。](#)

目標復旧時点 (RPO)

最後のデータリカバリポイントからの最大許容時間です。これにより、最後の回復時点からサービスが中断されるまでの間に許容できるデータ損失の程度が決まります。

目標復旧時間 (RTO)

サービスが中断から復旧までの最大許容遅延時間。

リファクタリング

[「7 Rs」を参照してください。](#)

リージョン

地理的エリア内の AWS リソースのコレクション。各 AWS リージョンは、耐障害性、安定性、耐障害性を提供するために、他のとは分離され、独立しています。詳細については、[AWS リージョン「を使用できるアカウントを指定する」](#)を参照してください。

回帰

数値を予測する機械学習手法。例えば、「この家はどれくらいの値段で売れるでしょうか?」という問題を解決するために、機械学習モデルは、線形回帰モデルを使用して、この家に関する既知の事実 (平方フィートなど) に基づいて家の販売価格を予測できます。

リHOST

[「7 R」を参照してください。](#)

リリース

デプロイプロセスで、変更を本番環境に昇格させること。

再配置

[「7 Rs」を参照してください。](#)

プラットフォーム変更

[「7 Rs」を参照してください。](#)

再購入

[「7 Rs」を参照してください。](#)

回復性

中断に耐えたり、中断から回復したりするアプリケーションの機能。で障害耐性を計画する場合、[高可用性とディザスタリカバリ](#)が一般的な考慮事項です AWS クラウド。詳細については、[AWS クラウド「レジリエンス」](#)を参照してください。

リソースベースのポリシー

Amazon S3 バケット、エンドポイント、暗号化キーなどのリソースにアタッチされたポリシー。このタイプのポリシーは、アクセスが許可されているプリンシパル、サポートされているアクション、その他の満たすべき条件を指定します。

実行責任者、説明責任者、協業先、報告先 (RACI) に基づくマトリックス

移行活動とクラウド運用に関わるすべての関係者の役割と責任を定義したマトリックス。マトリックスの名前は、マトリックスで定義されている責任の種類、すなわち責任 (R)、説明責任 (A)、協議 (C)、情報提供 (I) に由来します。サポート (S) タイプはオプションです。サポートを含めると、そのマトリックスは RASCI マトリックスと呼ばれ、サポートを除外すると RACI マトリックスと呼ばれます。

レスポンスコントロール

有害事象やセキュリティベースラインからの逸脱について、修復を促すように設計されたセキュリティコントロール。詳細については、Implementing security controls on AWSの[Responsive controls](#)を参照してください。

保持

[「7 Rs」を参照してください。](#)

廃止

[「7 R」を参照してください。](#)

ローテーション

攻撃者が認証情報にアクセスすることをより困難にするために、[シークレット](#)を定期的に更新するプロセス。

行と列のアクセス制御 (RCAC)

アクセスルールが定義された、基本的で柔軟な SQL 表現の使用。RCAC は行権限と列マスクで構成されています。

RPO

「目標[復旧時点](#)」を参照してください。

RTO

「目標[復旧時間](#)」を参照してください。

ランブック

特定のタスクを実行するために必要な手動または自動化された一連の手順。これらは通常、エラー率の高い反復操作や手順を合理化するために構築されています。

S

SAML 2.0

多くの ID プロバイダー (IdPs) が使用するオープンスタンダード。この機能により、フェデレーテッドシングルサインオン (SSO) が有効になるため、ユーザーは [AWS Management Console](#) したり [AWS CLI](#)、API オペレーションを呼び出したりできます。組織内のすべてのユーザーに対して IAM でユーザーを作成する必要はありません。SAML 2.0 ベースのフェデレーションの詳細については、IAM ドキュメントの[SAML 2.0 ベースのフェデレーションについて](#)を参照してください。

SCADA

「[監視コントロールとデータ収集](#)」を参照してください。

SCP

「[サービスコントロールポリシー](#)」を参照してください。

シークレット

では AWS Secrets Manager、暗号化された形式で保存するパスワードやユーザー認証情報などの機密情報または制限付き情報。シークレット値とそのメタデータで構成されます。シークレット値は、バイナリ、単一の文字列、または複数の文字列にすることができます。詳細については、[Secrets Manager ドキュメントの「Secrets Manager シークレットの内容」](#)を参照してください。

セキュリティコントロール

脅威アクターによるセキュリティ脆弱性の悪用を防止、検出、軽減するための、技術上または管理上のガードレール。セキュリティコントロールには、[予防的](#)、[検出的](#)、[???応答的](#)、[プロアクティブ](#) の4つの主なタイプがあります。

セキュリティ強化

アタックサーフェスを狭めて攻撃への耐性を高めるプロセス。このプロセスには、不要になったリソースの削除、最小特権を付与するセキュリティのベストプラクティスの実装、設定ファイル内の不要な機能の無効化、といったアクションが含まれています。

Security Information and Event Management (SIEM) システム

セキュリティ情報管理 (SIM) とセキュリティイベント管理 (SEM) のシステムを組み合わせたツールとサービス。SIEM システムは、サーバー、ネットワーク、デバイス、その他ソースからデータを収集、モニタリング、分析して、脅威やセキュリティ違反を検出し、アラートを発信します。

セキュリティレスポンスの自動化

セキュリティイベントに自動的に応答または修正するように設計された、事前定義されプログラムされたアクション。これらのオートメーションは、セキュリティのベストプラクティスを実装するのに役立つ検出的または[応答的な](#) AWS セキュリティコントロールとして機能します。自動レスポンスアクションの例としては、VPC セキュリティグループの変更、Amazon EC2 インスタンスへのパッチ適用、認証情報のローテーションなどがあります。

サーバー側の暗号化

送信先にあるデータの、それを受け取る AWS のサービス による暗号化。

サービスコントロールポリシー (SCP)

AWS Organizationsの組織内の、すべてのアカウントのアクセス許可を一元的に管理するポリシー。SCP は、管理者がユーザーまたはロールに委任するアクションに、ガードレールを定義したり、アクションの制限を設定したりします。SCP は、許可リストまたは拒否リストとして、許可または禁止するサービスやアクションを指定する際に使用できます。詳細については、AWS Organizations ドキュメントの「[サービスコントロールポリシー](#)」を参照してください。

サービスエンドポイント

のエンドポイントの URL AWS のサービス。ターゲットサービスにプログラムで接続するには、エンドポイントを使用します。詳細については、AWS 全般のリファレンスの「[AWS のサービス エンドポイント](#)」を参照してください。

サービスレベルアグリーメント (SLA)

サービスのアップタイムやパフォーマンスなど、IT チームがお客様に提供すると約束したものを明示した合意書。

サービスレベルインジケータ (SLI)

エラー率、可用性、スループットなど、サービスのパフォーマンス側面の測定。

サービスレベルの目標 (SLO)

サービスレベルのインジケータによって測定される、サービスの状態を表すターゲットメトリクス。

責任共有モデル

クラウドのセキュリティとコンプライアンス AWS についてと共有する責任を説明するモデル。AWS はクラウドのセキュリティを担当しますが、お客様はクラウドのセキュリティを担当します。詳細については、責任共有モデルを参照してください。

SIEM

「セキュリティ情報とイベント管理システム」を参照してください。

単一障害点 (SPOF)

システムを中断する可能性のあるアプリケーションの単一の重要なコンポーネントの障害。

SLA

「サービスレベルアグリーメント」を参照してください。

SLI

「サービスレベルインジケータ」を参照してください。

SLO

「サービスレベルの目標」を参照してください。

split-and-seed モデル

モダナイゼーションプロジェクトのスケーリングと加速のためのパターン。新機能と製品リリースが定義されると、コアチームは解放されて新しい製品チームを作成します。これにより、お客様の組織の能力とサービスの拡張、デベロッパーの生産性の向上、迅速なイノベーションのサ

ポートに役立ちます。詳細については、[「」の「アプリケーションをモダナイズするための段階的アプローチ AWS クラウド」](#)を参照してください。

SPOF

[単一障害点](#)を参照してください。

star スキーマ

トランザクションデータまたは測定データを保存するために1つの大きなファクトテーブルを使用し、データ属性を保存するために1つ以上の小さなディメンションテーブルを使用するデータベースの組織構造。この構造は、[データウェアハウス](#)またはビジネスインテリジェンスの目的で使用するよう設計されています。

strangler fig パターン

レガシーシステムが廃止されるまで、システム機能を段階的に書き換えて置き換えることにより、モノリシックシステムをモダナイズするアプローチ。このパターンは、宿主の樹木から根を成長させ、最終的にその宿主を包み込み、宿主に取って代わるイチジクのつるを例えています。そのパターンは、モノリシックシステムを書き換えるときのリスクを管理する方法として [Martin Fowler](#) により提唱されました。このパターンの適用方法の例については、[コンテナと Amazon API Gateway](#) を使用して、従来の Microsoft ASP.NET (ASMX) ウェブサービスを段階的にモダナイズを参照してください。

サブネット

VPC 内の IP アドレスの範囲。サブネットは、1つのアベイラビリティゾーンに存在する必要があります。

監視統制とデータ収集 (SCADA)

製造では、ハードウェアとソフトウェアを使用して物理アセットと生産オペレーションをモニタリングするシステム。

対称暗号化

データの暗号化と復号に同じキーを使用する暗号化のアルゴリズム。

合成テスト

ユーザーインタラクションをシミュレートして潜在的な問題を検出したり、パフォーマンスをモニタリングしたりする方法でシステムをテストします。[Amazon CloudWatch Synthetics](#) を使用してこれらのテストを作成できます。

T

タグ

AWS リソースを整理するためのメタデータとして機能するキーと値のペア。タグは、リソースの管理、識別、整理、検索、フィルタリングに役立ちます。詳細については、「[AWS リソースのタグ付け](#)」を参照してください。

ターゲット変数

監督された機械学習でお客様が予測しようとしている値。これは、結果変数のことも指します。例えば、製造設定では、ターゲット変数が製品の欠陥である可能性があります。

タスクリスト

ランブックの進行状況を追跡するために使用されるツール。タスクリストには、ランブックの概要と完了する必要がある一般的なタスクのリストが含まれています。各一般的なタスクには、推定所要時間、所有者、進捗状況が含まれています。

テスト環境

[「環境」](#)を参照してください。

トレーニング

お客様の機械学習モデルに学習するデータを提供すること。トレーニングデータには正しい答えが含まれている必要があります。学習アルゴリズムは入力データ属性をターゲット (お客様が予測したい答え) にマッピングするトレーニングデータのパターンを検出します。これらのパターンをキャプチャする機械学習モデルを出力します。そして、お客様が機械学習モデルを使用して、ターゲットがわからない新しいデータでターゲットを予測できます。

トランジットゲートウェイ

VPC とオンプレミスネットワークを相互接続するために使用できる、ネットワークの中継ハブ。詳細については、AWS Transit Gateway ドキュメントの「[トランジットゲートウェイとは](#)」を参照してください。

トランクベースのワークフロー

デベロッパーが機能ブランチで機能をローカルにビルドしてテストし、その変更をメインブランチにマージするアプローチ。メインブランチはその後、開発環境、本番前環境、本番環境に合わせて順次構築されます。

信頼されたアクセス

ユーザーに代わって AWS Organizations およびそのアカウントで組織内でタスクを実行するために指定するサービスへのアクセス許可を付与します。信頼されたサービスは、サービスにリンクされたロールを必要なときに各アカウントに作成し、ユーザーに代わって管理タスクを実行します。詳細については、ドキュメント [AWS Organizations の「を他の AWS のサービスで使用する AWS Organizations」](#) を参照してください。

チューニング

機械学習モデルの精度を向上させるために、お客様のトレーニングプロセスの側面を変更する。例えば、お客様が機械学習モデルをトレーニングするには、ラベル付けセットを生成し、ラベルを追加します。これらのステップを、異なる設定で複数回繰り返して、モデルを最適化します。

ツーピザチーム

2つのピザを食べることができる小さな DevOps チーム。ツーピザチームの規模では、ソフトウェア開発におけるコラボレーションに最適な機会が確保されます。

U

不確実性

予測機械学習モデルの信頼性を損なう可能性がある、不正確、不完全、または未知の情報を指す概念。不確実性には、次の2つのタイプがあります。認識論的不確実性は、限られた、不完全なデータによって引き起こされ、弁論的不確実性は、データに固有のノイズとランダム性によって引き起こされます。詳細については、[深層学習システムにおける不確実性の定量化](#) ガイドを参照してください。

未分化なタスク

ヘビーリフティングとも呼ばれ、アプリケーションの作成と運用には必要だが、エンドユーザーに直接的な価値をもたらさなかったり、競争上の優位性をもたらしたりしない作業です。未分化なタスクの例としては、調達、メンテナンス、キャパシティプランニングなどがあります。

上位環境

[「環境」](#) を参照してください。

V

バキューミング

ストレージを再利用してパフォーマンスを向上させるために、増分更新後にクリーンアップを行うデータベースのメンテナンス操作。

バージョンコントロール

リポジトリ内のソースコードへの変更など、変更を追跡するプロセスとツール。

VPC ピアリング

プライベート IP アドレスを使用してトラフィックをルーティングできる、2 つの VPC 間の接続。詳細については、Amazon VPC ドキュメントの「[VPC ピア機能とは](#)」を参照してください。

脆弱性

システムのセキュリティを脅かすソフトウェアまたはハードウェアの欠陥。

W

ウォームキャッシュ

頻繁にアクセスされる最新の関連データを含むバッファキャッシュ。データベースインスタンスはバッファキャッシュから、メインメモリまたはディスクからよりも短い時間で読み取りを行うことができます。

ウォームデータ

アクセス頻度の低いデータ。この種類のデータをクエリする場合、通常は適度に遅いクエリでも問題ありません。

ウィンドウ関数

現在のレコードに関連する行のグループに対して計算を実行する SQL 関数。ウィンドウ関数は、移動平均の計算や、現在の行の相対位置に基づく行の値へのアクセスなどのタスクの処理に役立ちます。

ワークロード

ビジネス価値をもたらすリソースとコード (顧客向けアプリケーションやバックエンドプロセスなど) の総称。

ワークストリーム

特定のタスクセットを担当する移行プロジェクト内の機能グループ。各ワークストリームは独立していますが、プロジェクト内の他のワークストリームをサポートしています。たとえば、ポートフォリオワークストリームは、アプリケーションの優先順位付け、ウェーブ計画、および移行メタデータの収集を担当します。ポートフォリオワークストリームは、これらの設備を移行ワークストリームで実現し、サーバーとアプリケーションを移行します。

WORM

[「書き込み 1 回」](#)を参照し、[多くの を読み取ります](#)。

WQF

[「AWS ワークロード認定フレームワーク」](#)を参照してください。

Write Once, Read Many (WORM)

データを 1 回書き込み、データの削除や変更を防ぐストレージモデル。承認されたユーザーは、必要な回数だけデータを読み取ることができますが、変更することはできません。このデータストレージインフラストラクチャは [イミュータブルな](#) と見なされます。

Z

ゼロデイエクスプロイト

[ゼロデイ脆弱性](#) を利用する攻撃、通常はマルウェア。

ゼロデイ脆弱性

実稼働システムにおける未解決の欠陥または脆弱性。脅威アクターは、このような脆弱性を利用してシステムを攻撃する可能性があります。開発者は、よく攻撃の結果で脆弱性に気付きます。

ゾンビアプリケーション

平均 CPU およびメモリ使用率が 5% 未満のアプリケーション。移行プロジェクトでは、これらのアプリケーションを廃止するのが一般的です。

翻訳は機械翻訳により提供されています。提供された翻訳内容と英語版の間で齟齬、不一致または矛盾がある場合、英語版が優先します。