aws

Guida per gli sviluppatori

Amazon Managed Streaming per Apache Kafka

Amazon Managed Streaming per Apache Kafka: Guida per gli sviluppatori

Table of Contents

Benvenuto nella Guida per gli sviluppatori di Amazon MSK

Ti diamo il benvenuto nella Guida per gli sviluppatori di Amazon MSK. I seguenti argomenti possono aiutarti a iniziare a usare questa guida, in base a ciò che stai cercando di fare.

- Crea un cluster Amazon MSK seguendo il tutorial Guida introduttiva all'utilizzo di Amazon MSK.
- Per approfondire le funzionalità di Amazon MSK, consulta la sezione Amazon MSK: come funziona.
- Esegui Apache Kafka senza dover gestire e dimensionare la capacità del cluster con MSK Serverless.
- Utilizza MSK Connect per lo streaming di dati da e verso il cluster Apache Kafka.
- Replicatore MSKUtilizzalo per replicare in modo affidabile i dati tra cluster Amazon MSK in AWS regioni diverse o uguali.

Per i dettagli, le funzionalità principali e i prezzi del prodotto, consulta la pagina del servizio Amazon MSK.

Che cos'è Amazon MSK?

Amazon Managed Streaming for Apache Kafka (Amazon MSK) è un servizio completamente gestito che consente di costruire ed eseguire applicazioni che utilizzano Apache Kafka per elaborare i dati in streaming. Amazon MSK fornisce le operazioni del piano di controllo, ad esempio quelle per la creazione, l'aggiornamento e l'eliminazione di cluster. Consente di utilizzare operazioni del piano dati Apache Kafka, come quelle per la produzione e il consumo di dati. Esegue versioni open-source di Apache Kafka. Ciò significa che le applicazioni, gli strumenti e i plugin esistenti dei partner e della comunità Apache Kafka sono supportati senza richiedere modifiche al codice dell'applicazione. Puoi utilizzare Amazon MSK per creare cluster che utilizzano le versioni di Apache Kafka elencate nella sezione the section called “Versioni di Apache Kafka supportate”.

Questi componenti descrivono l'architettura di Amazon MSK:

- Nodi dei broker: quando crei un cluster Amazon MSK, specifichi quanti nodi dei broker desideri che Amazon MSK crei in ciascuna zona di disponibilità. Il minimo è un broker per zona di disponibilità. Ogni zona di disponibilità dispone di una propria sottorete VPC.
- ZooKeeper nodi: Amazon MSK crea anche i ZooKeeper nodi Apache per te. Apache ZooKeeper è un server open source che consente un coordinamento distribuito altamente affidabile.

- Controller KRAFT: la comunità Apache Kafka ha sviluppato KRAFT per sostituire Apache per la gestione dei metadati nei cluster Apache Kafka. ZooKeeper In modalità KRAFT, i metadati del cluster vengono propagati all'interno di un gruppo di controller Kafka, che fanno parte del cluster Kafka, anziché tra i nodi. ZooKeeper I controller Kraft sono inclusi senza costi aggiuntivi per l'utente e non richiedono alcuna configurazione o gestione aggiuntiva da parte dell'utente.

  > **Note**
  >
  > Dalla versione 3.7.x di Apache Kafka su MSK, è possibile creare cluster che utilizzano la modalità KRAFT anziché la modalità. ZooKeeper

- Produttori, consumatori e creatori di argomenti: Amazon MSK ti consente di utilizzare le operazioni sul piano dati di Apache Kafka per creare argomenti e produrre e utilizzare dati.
- Operazioni del cluster È possibile utilizzare le AWS Management Console, the AWS Command Line Interface (AWS CLI) o le API nell'SDK per eseguire operazioni sul piano di controllo. Ad esempio, puoi creare o eliminare un cluster Amazon MSK, elencare tutti i cluster di un account, visualizzare le proprietà di un cluster e aggiornare il numero e il tipo di broker in un cluster.

Amazon MSK riconosce gli scenari di errore più comuni e avvia automaticamente il ripristino per cluster per permettere alle applicazioni produttore e consumatore di continuare le operazioni di scrittura e lettura con impatto minimo. Quando Amazon MSK rileva un errore del broker, attenua l'errore o sostituisce il broker non integro o non raggiungibile con uno nuovo. Inoltre, ove possibile, riutilizza lo storage del broker precedente per ridurre i dati che devono essere replicati da Apache Kafka. L'impatto sulla disponibilità è limitato al tempo richiesto da Amazon MSK per completare il rilevamento e il ripristino. Dopo un ripristino, le applicazioni produttore e consumatore possono continuare a comunicare con gli stessi indirizzi IP del broker utilizzati prima dell'errore.

Configurazione di Amazon MSK

Prima di utilizzare Amazon MSK per la prima volta, è necessario completare le seguenti operazioni.

Attività

- Registrati per AWS
- Esecuzione del download di librerie e strumenti

Registrati per AWS

Quando ti registri AWS, il tuo account Amazon Web Services viene automaticamente registrato per tutti i servizi in AWS, incluso Amazon MSK. Ti vengono addebitati solo i servizi che utilizzi.

Se hai già un AWS account, passa all'attività successiva. Se non disponi di un account AWS , utilizza la seguente procedura per crearne uno.

Registrazione per un account Amazon Web Services

1. Apri la pagina all'indirizzo https://portal.aws.amazon.com/billing/signup.
2. Segui le istruzioni online.

   Nel corso della procedura di registrazione riceverai una telefonata, durante la quale sarà necessario inserire un codice di verifica attraverso la tastiera del telefono.

   Quando ti iscrivi a un Account AWS, ne Utente root dell'account AWSviene creato un. L'utente root dispone dell'accesso a tutte le risorse e tutti i Servizi AWS nell'account. Come procedura consigliata in materia di sicurezza, assegnate l'accesso amministrativo a un utente e utilizzate solo l'utente root per eseguire attività che richiedono l'accesso da parte dell'utente root.

Esecuzione del download di librerie e strumenti

Le librerie e gli strumenti seguenti semplificano l'utilizzo di Amazon MSK:

- L'AWS Command Line Interface (AWS CLI) supporta Amazon MSK. Ti AWS CLI consente di controllare più Amazon Web Services dalla riga di comando e di automatizzarli tramite script. Effettua AWS CLI l'upgrade alla versione più recente per assicurarti che supporti le funzionalità di Amazon MSK documentate in questa guida per l'utente. Per istruzioni dettagliate su come

aggiornare la AWS CLI, consulta la pagina Installing the AWS Command Line Interface. Dopo aver installato AWS CLI, devi configurarlo. Per informazioni su come configurare AWS CLI, consulta aws configure.

- La Documentazione di riferimento a Streaming gestito da Amazon per Apache Kafka documenta le operazioni dell'API supportate da Amazon MSK.
- Gli SDK Amazon Web Services per Go, Java, .NET JavaScript, Node.js, PHP, Python e Ruby includono il supporto e gli esempi di Amazon MSK.

Guida introduttiva all'utilizzo di Amazon MSK

In questa sezione viene illustrato un esempio di come creare un cluster MSK, produrre e utilizzare dati, nonché monitorare l'integrità del cluster utilizzando i parametri. Questo esempio non rappresenta tutte le opzioni che è possibile scegliere quando si crea un cluster MSK. In diverse parti di questo tutorial verranno scelte opzioni predefinite per semplicità. Ciò non significa che siano le uniche opzioni che funzionano per la configurazione di un cluster MSK o delle istanze client.

Argomenti

Passaggio 1: creazione di un cluster Amazon MSK

In questo passaggio della Guida introduttiva all'utilizzo di Amazon MSK, crei un cluster Amazon MSK.

Per creare un cluster Amazon MSK utilizzando AWS Management Console

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Scegli Create cluster (Crea cluster).
3. Per Metodo di creazione, lascia selezionata l'opzione Creazione rapida. L'opzione Creazione rapida consente di creare un cluster con le impostazioni predefinite.
4. Per Nome cluster, inserisci un nome descrittivo per il cluster. Ad esempio, **`MSKTutorialCluster`**.
5. In Proprietà generali del cluster, scegli Assegnato come Tipo di cluster.
6. Dalla tabella in Tutte le impostazioni del cluster, copia i valori delle seguenti impostazioni e salvali perché ti serviranno più avanti in questo tutorial:

- VPC
- Sottoreti
- Gruppi di sicurezza associati al VPC

7. Scegli Create cluster (Crea cluster).
8. Verifica lo Stato del cluster nella pagina Riepilogo del cluster. Quando Amazon MSK assegna il cluster, lo stato passa da Creazione in corso ad Attivo. Quando lo stato è Attivo, puoi connetterti al cluster. Per ulteriori informazioni sugli stati del cluster, consulta la pagina Stati dei cluster.

Fase successiva

Fase 2: creazione di un ruolo IAM

Fase 2: creazione di un ruolo IAM

In questo passaggio, eseguirai due attività. La prima attività consiste nel creare una policy IAM che consenta l'accesso alla creazione di argomenti nel cluster e all'invio di dati a tali argomenti. La seconda attività consiste nel creare un ruolo IAM e associarvi questa policy. In un passaggio successivo, si crea un computer client che assume questo ruolo e lo utilizza per creare un argomento nel cluster e per inviare dati a quell'argomento.

Creazione di una policy IAM che consenta di creare argomenti e scrivere su di essi

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro di navigazione, seleziona Policy.
3. Scegliere Create Policy (Crea policy).
4. Scegli la scheda JSON, quindi sostituisci il JSON nella finestra dell'editor con il seguente JSON.

   Sostituisci la *regione* con il codice della AWS regione in cui hai creato il cluster. Sostituisci *Account-ID* con il tuo ID account. Sostituisci *MSK TutorialCluster* con il nome del cluster.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:Connect",
```

```
                "kafka-cluster:AlterCluster",
                "kafka-cluster:DescribeCluster"
            ],
            "Resource": [
                "arn:aws:kafka:region:Account-ID:cluster/MSKTutorialCluster/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:*Topic*",
                "kafka-cluster:WriteData",
                "kafka-cluster:ReadData"
            ],
            "Resource": [
                "arn:aws:kafka:region:Account-ID:topic/MSKTutorialCluster/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:AlterGroup",
                "kafka-cluster:DescribeGroup"
            ],
            "Resource": [
                "arn:aws:kafka:region:Account-ID:group/MSKTutorialCluster/*"
            ]
        }
    ]
}
```

Per ricevere istruzioni su come scrivere policy sicure, consulta la pagina the section called “Controllo degli accessi IAM”.

5. Scegliere Next: Tags (Successivo: Tag).
6. Scegliere Next:Review (Successivo: Rivedi).
7. Per il nome della policy, inserisci un nome descrittivo, ad esempio msk-tutorial-policy.
8. Scegli Crea policy.

Creazione di un ruolo IAM e collegamento della policy al ruolo

1. Nel riquadro di navigazione, seleziona Ruoli.

2. Scegli Crea ruolo.
3. In Casi di utilizzo comuni, scegli EC2, quindi scegli Successivo: autorizzazioni.
4. Nella casella di ricerca, inserisci il nome della policy creata in precedenza per questo tutorial. Seleziona quindi la casella a sinistra della policy.
5. Scegliere Next: Tags (Successivo: Tag).
6. Scegliere Next:Review (Successivo: Rivedi).
7. Per il nome del ruolo, inserisci un nome descrittivo, ad esempio msk-tutorial-role.
8. Scegli Crea ruolo.

Fase successiva

Passaggio 3: creazione di un computer client

Passaggio 3: creazione di un computer client

In questo passaggio della Guida introduttiva all'utilizzo di Amazon MSK, crei un computer client. Utilizza questo computer client per creare un argomento che produce e consuma dati. Per semplicità, creerai questo computer client nel VPC associato al cluster MSK in modo che il client possa connettersi facilmente al cluster.

Per creare un computer client

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Scegliere Launch Instances (Avvia istanze).
3. Inserisci un Nome per il computer client, ad esempio **`MSKTutorialClient`**.
4. Lascia Amazon Linux 2 AMI (HVM) - Kernel 5.10, tipo di volume SSD selezionato per Tipo di Amazon Machine Image (AMI).
5. Lascia selezionato il tipo di istanza t2.micro.
6. In Coppia di chiavi (accesso), scegli Crea una nuova coppia di chiavi. Inserisci **`MSKKeyPair`** in Nome della coppia di chiavi, quindi scegli Scarica coppia di chiavi. In alternativa, è possibile utilizzare una coppia di chiavi esistente.
7. Espandi la sezione Dettagli avanzati e scegli il ruolo IAM che hai creato nel Passaggio 2: creazione di un ruolo IAM.

8. Scegliere Launch Instance (Avvia istanza).
9. Scegliere View Instances (Vedi istanze). Quindi, nella colonna Gruppi di sicurezza, scegli il gruppo di sicurezza associato alla nuova istanza. Copia l'ID del gruppo di sicurezza e salvalo per un secondo momento.
10. Accedi alla console Amazon VPC all'indirizzo https://console.aws.amazon.com/vpc/.
11. Fai clic su Security Groups (Gruppi di sicurezza) nel riquadro di navigazione. Trova il gruppo di sicurezza del quale hai salvato l'ID in the section called “Fase 1: creazione di un cluster”.
12. Nella scheda Regole in entrata, scegli Modifica le regole in entrata.
13. Scegli Aggiungi regola.
14. Nella nuova regola, scegliere All traffic (Tutto il traffico) nella colonna Type (Tipo) . Nel secondo campo della colonna Origine, inserisci l'ID del gruppo di sicurezza del computer client. Questo è il gruppo il cui nome hai salvato dopo aver avviato l'istanza del computer client.
15. Scegliere Salva regole. Ora il gruppo di sicurezza del cluster può accettare il traffico proveniente dal gruppo di sicurezza del computer client.

Fase successiva

Passaggio 4: creazione di un argomento

Passaggio 4: creazione di un argomento

In questo passaggio della Guida introduttiva all'utilizzo di Amazon MSK, installi librerie e strumenti client di Apache Kafka sul computer client e quindi crei un argomento.

> **Warning**
>
> I numeri di versione di Apache Kafka utilizzati in questo tutorial sono solo degli esempi. Ti consigliamo di utilizzare la stessa versione del client della versione del cluster MSK. In una versione precedente del client potrebbero mancare alcune funzionalità e correzioni di bug critici.

Individuazione della versione del cluster MSK

1. Vai a https://eu-west-2.console.aws.amazon.com/msk/

2. Seleziona il cluster MSK.
3. Prendi nota della versione di Apache Kafka utilizzata nel cluster.
4. Sostituisci le istanze dei numeri di versione di Amazon MSK in questo tutorial con la versione ottenuta nel passaggio 3.

Per creare un argomento sul computer client

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, seleziona Istanze. Quindi seleziona la casella di controllo accanto al nome del computer client che hai creato in Passaggio 3: creazione di un computer client.
3. Scegliere Actions (Operazioni), quindi selezionare Connect (Connetti). Segui le istruzioni riportate nella console per connetterti al computer client.
4. Installare Java sul computer client eseguendo il seguente comando:

```
sudo yum -y install java-11
```

5. Eseguire il seguente comando per scaricare Apache Kafka.

```
wget https://archive.apache.org/dist/kafka/{YOUR MSK VERSION}/kafka_2.13-{YOUR MSK
 VERSION}.tgz
```

> **Note**
>
> Se desideri utilizzare un sito mirror diverso da quello utilizzato in questo comando, puoi sceglierne uno diverso sul sito Web di Apache .

6. Eseguire il comando seguente nella directory in cui è stato scaricato il file TAR nella fase precedente.

```
tar -xzf kafka_2.13-{YOUR MSK VERSION}.tgz
```

7. Vai alla directory `kafka_2.13-{YOUR MSK VERSION}/libs`, quindi esegui il comando per scaricare il file JAR IAM di Amazon MSK. Il file JAR IAM di Amazon MSK consente al computer client di accedere al cluster.

```
wget https://github.com/aws/aws-msk-iam-auth/releases/download/v1.1.1/aws-msk-iam-
auth-1.1.1-all.jar
```

8. Vai alla directory `kafka_2.13-{YOUR MSK VERSION}/bin`. Copia le impostazioni delle proprietà seguenti e incollale in un nuovo file. Assegna al file il nome **`client.properties`** e salvalo.

```
security.protocol=SASL_SSL
sasl.mechanism=AWS_MSK_IAM
sasl.jaas.config=software.amazon.msk.auth.iam.IAMLoginModule required;
sasl.client.callback.handler.class=software.amazon.msk.auth.iam.IAMClientCallbackHandler
```

9. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
10. Attendi che lo stato del cluster diventi Attivo. Questo processo potrebbe richiedere diversi minuti. Dopo che lo stato diventa Attivo, scegli il nome del cluster. Verrà visualizzata una pagina contenente il riepilogo del cluster.
11. Scegli Visualizza le informazioni sul client.
12. Copia la stringa di connessione per l'endpoint privato.

    Otterrai tre endpoint per ciascuno dei broker. È richiesto un solo endpoint del broker per il passaggio successivo.
13. Esegui il comando seguente, sostituendo *`BootstrapServerString`* con uno degli endpoint del broker ottenuti nel passaggio precedente.

```
<path-to-your-kafka-installation>/bin/kafka-topics.sh --create --bootstrap-server
 BootstrapServerString --command-config client.properties --replication-factor 3 --
partitions 1 --topic MSKTutorialTopic
```

    Se il comando va a buon fine, viene visualizzato il seguente messaggio: `Created topic MSKTutorialTopic.`

Fase successiva

Passaggio 5: produzione e utilizzo di dati

Passaggio 5: produzione e utilizzo di dati

In questo passaggio della Guida introduttiva all'utilizzo di Amazon MSK, produci e utilizzi dati.

Per produrre e consumare messaggi

1. Eseguire il comando seguente per avviare un produttore della console. Sostituisci *BootstrapServerString* con la stringa di connessione in testo semplice ottenuta in Crea un argomento. Per istruzioni su come recuperare questa stringa di connessione, consulta la sezione Recupero dei broker di bootstrap per un cluster Amazon MSK.

   ```
   <path-to-your-kafka-installation>/bin/kafka-console-producer.sh --
   broker-list BootstrapServerString --producer.config client.properties --
   topic MSKTutorialTopic
   ```

2. Immettere qualsiasi messaggio desiderato e premere Enter (Invio). Ripetere questa fase due o tre volte. Ogni volta che si immette una riga e si preme Enter (Invio), tale riga viene inviata al cluster Apache Kafka come un messaggio separato.
3. Mantenere aperta la connessione al computer client, quindi aprire una seconda connessione separata al computer in una nuova finestra.
4. Nel comando seguente, sostituisci *BootstrapServerString con la stringa* di connessione in testo semplice salvata in precedenza. Quindi, per creare un utente della console, esegui il comando seguente con la tua seconda connessione al computer client.

   ```
   <path-to-your-kafka-installation>/bin/kafka-console-consumer.sh --bootstrap-
   server BootstrapServerString --consumer.config client.properties --
   topic MSKTutorialTopic --from-beginning
   ```

   Si iniziano a vedere i messaggi immessi in precedenza quando è stato utilizzato il comando produttore della console.
5. Immettere altri messaggi nella finestra del produttore e guardali apparire nella finestra del consumatore.

Fase successiva

Fase 6: Usa Amazon CloudWatch per visualizzare i parametri di Amazon MSK

Fase 6: Usa Amazon CloudWatch per visualizzare i parametri di Amazon MSK

In questa fase di Guida introduttiva all'uso di Amazon MSK, esamini i parametri di Amazon MSK in Amazon. CloudWatch

Per visualizzare i parametri di Amazon MSK in CloudWatch

1. Apri la CloudWatch console all'indirizzo https://console.aws.amazon.com/cloudwatch/.
2. Nel riquadro di navigazione, seleziona Parametri.
3. Scegliere la scheda All metrics (Tutti i parametri), quindi selezionare AWS/Kafka.
4. Per visualizzare i parametri a livello di broker, scegliere Broker ID, Cluster Name (ID broker, nome cluster). Per i parametri a livello di cluster, scegliere Cluster Name (Nome cluster).
5. (Facoltativo) Nel riquadro grafico, selezionate una statistica e un periodo di tempo, quindi create un CloudWatch allarme utilizzando queste impostazioni.

Fase successiva

Passaggio 7: Eliminare le AWS risorse create per questo tutorial

Passaggio 7: Eliminare le AWS risorse create per questo tutorial

Nel passaggio finale della Guida introduttiva all'utilizzo di Amazon MSK, elimini il cluster MSK e il computer client che hai creato per questo tutorial.

Per eliminare le risorse utilizzando il AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli il nome del cluster. Ad esempio, MSK. TutorialCluster
3. Selezionare Actions (Operazioni), quindi selezionare Delete (Elimina).
4. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
5. Scegli l'istanza che hai creato per il computer client, ad esempio **`MSKTutorialClient`**.
6. Scegli Stato istanza, quindi scegli Termina istanza.

Eliminazione del ruolo e della policy IAM

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro di navigazione, seleziona Ruoli.
3. Nella casella di ricerca, inserisci il nome del ruolo IAM creato per questo tutorial.
4. Seleziona il ruolo. Quindi scegli Elimina ruolo e conferma l'eliminazione.
5. Nel riquadro di navigazione, seleziona Policy.
6. Nella casella di ricerca, inserisci il nome della policy creata per questo tutorial.
7. Scegli la policy per aprirne la pagina di riepilogo. Nella pagina di Riepilogo della policy, seleziona Elimina policy.
8. Scegli Elimina.

Amazon MSK: come funziona

Un cluster Amazon MSK è la risorsa Amazon MSK primaria che puoi creare nel tuo account. Negli argomenti di questa sezione viene descritto come eseguire le operazioni comuni di Amazon MSK. Per un elenco di tutte le operazioni che puoi eseguire su un cluster MSK, consulta le risorse seguenti:

- Il AWS Management Console
- La documentazione di riferimento all'API di Amazon MSK
- La documentazione di riferimento ai comandi della CLI di Amazon MSK

Argomenti

Creazione di un cluster Amazon MSK

> **Important**
>
> Non è possibile modificare il VPC di un cluster Amazon MSK dopo aver creato il cluster.

Prima di poter creare un cluster Amazon MSK, è necessario aver configurato un Amazon Virtual Private Cloud (VPC) e delle sottoreti al suo interno.

Sono necessarie due sottoreti in due zone di disponibilità diverse nella regione Stati Uniti occidentali (California settentrionale). Per altre regioni in cui è disponibile Amazon MSK, è possibile specificare due o tre sottoreti. Le sottoreti devono trovarsi tutte in zone di disponibilità differenti. Quando crei un cluster, Amazon MSK distribuisce i nodi dei broker in modo uniforme nelle sottoreti specificate.

Dimensioni dei broker

Quando crei un cluster Amazon MSK, specifichi la dimensione dei broker che desideri che abbia. Amazon MSK supporta le seguenti dimensioni di broker:

- kafka.t3.small
- kafka.m5.large, kafka.m5.xlarge, kafka.m5.2xlarge, kafka.m5.4xlarge, kafka.m5.8xlarge, kafka.m5.12xlarge, kafka.m5.16xlarge, kafka.m5.24xlarge
- kafka.m7g.large, kafka.m7g.xlarge, kafka.m7g.2xlarge, kafka.m7g.4xlarge, kafka.m7g.8xlarge, kafka.m7g.12xlarge, kafka.m7g.12xlarge, kafka.m7g.16xlarge

I broker M7g utilizzano processori AWS Graviton (processori personalizzati basati su ARM creati da Amazon Web Services). I broker M7g offrono un rapporto prezzo/prestazioni migliorato rispetto alle istanze M5 comparabili. I broker M7g consumano meno energia rispetto alle istanze M5 comparabili.

I broker M7g Graviton non sono disponibili nelle seguenti regioni: CDG (Parigi), CGK (Giacarta), CPT (Città del Capo), DXB (Dubai), HKG (Hong Kong), KIX (Osaka), LHR (Londra), MEL (Melbourne), MXP (Milano), OSU (Stati Uniti orientali), PDT (Stati Uniti occidentali), TLV (Tel Aviv), YY YC (Calgary), ZRH (Zurigo).

MSK supporta i broker M7g su cluster che eseguono una delle seguenti versioni di Kafka:

- 2.8.2. A più livelli
- 3.3.2
- 3.4.0
- 3.5.1
- 3.6.0 con storage su più livelli
- 3.7.x
- 3.7.x.kraft

I broker M7g e M5 offrono prestazioni di throughput di base più elevate rispetto ai broker T3 e sono consigliati per carichi di lavoro di produzione. I broker M7g e M5 possono anche avere più partizioni per broker rispetto ai broker T3. Usa i broker M7g o M5 se esegui carichi di lavoro di livello di produzione più grandi o richiedi un numero maggiore di partizioni. Per ulteriori informazioni sulle dimensioni delle istanze M7g e M5, consulta Amazon EC2 General Purpose Instances.

I broker T3 possono utilizzare i crediti della CPU per incrementare temporaneamente le prestazioni. Utilizza i broker T3 per lo sviluppo a basso costo, se stai provando carichi di lavoro di streaming di piccole e medie dimensioni o se disponi di carichi di lavoro di streaming a throughput basso con picchi temporanei di throughput. Ti consigliamo di eseguire un proof-of-concept test per determinare se i broker T3 sono sufficienti per la produzione o per un carico di lavoro critico. Per ulteriori informazioni sulle dimensioni dei broker T3, consulta Amazon EC2T3 Instances.

Per ulteriori informazioni su come scegliere le dimensioni dei broker, consulta. Best practice

Creazione di un cluster utilizzando AWS Management Console

Questo processo descrive l'attività comune di creazione di un cluster predisposto utilizzando opzioni di creazione personalizzate. È possibile selezionare altre opzioni nella console MSK per creare un cluster serverless.

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli Create cluster (Crea cluster).
3. Per il metodo di creazione del cluster, scegli Creazione personalizzata.
4. Specificate un nome del cluster che sia unico e che non superi i 64 caratteri.
5. Per Tipo di cluster, scegli Provisioned, che consente di specificare il numero di broker, le dimensioni del broker e la capacità di archiviazione del cluster.
6. Seleziona la versione di Apache Kafka che desideri eseguire sui broker. Per visualizzare un confronto tra le funzionalità di MSK supportate da ciascuna versione di Apache Kafka, seleziona Visualizza compatibilità tra le versioni.
7. A seconda della versione di Apache Kafka selezionata, potresti avere la possibilità di scegliere la modalità Metadati del cluster: o KRAFT. ZooKeeper
8. Seleziona la dimensione del broker da utilizzare per il cluster in base alle esigenze di calcolo, memoria e archiviazione del cluster. Consultare ???.
9. Seleziona il numero di zone in cui sono distribuiti i broker.

10. Specificate il numero di broker che desiderate che MSK crei in ogni zona di disponibilità. Il minimo è un broker per zona di disponibilità e il massimo è 30 broker per cluster per i cluster ZooKeeper basati su Kraft e 60 broker per cluster.
11. Seleziona la quantità iniziale di storage che desideri assegnare al tuo cluster. Non è possibile ridurre la capacità di archiviazione dopo aver creato il cluster.
12. A seconda della dimensione del broker (dimensione dell'istanza) selezionata, è possibile specificare il throughput di storage Provisioned per broker. Per abilitare questa opzione, scegli la dimensione del broker (dimensione dell'istanza) kafka.m5.4xlarge o maggiore per x86 e kafka.m7g.2xlarge o maggiore per le istanze basate su Graviton. Per informazioni, consulta ???.
13. Seleziona un'opzione in modalità di archiviazione cluster, solo storage EBS o storage su più livelli e storage EBS.
14. Se desideri creare e utilizzare una configurazione cluster personalizzata (o se hai già una configurazione del cluster salvata), scegli una configurazione. Altrimenti, puoi creare il cluster utilizzando la configurazione cluster predefinita di Amazon MSK. Per informazioni sulle configurazioni di Amazon MSK, consulta la sezione Configurazione.
15. Seleziona Avanti.
16. Per le impostazioni di rete, scegli il VPC che desideri utilizzare per il cluster.
17. In base al numero di zone selezionato in precedenza, specifica le zone di disponibilità e le sottoreti in cui verranno implementate i broker. Le sottoreti devono trovarsi in zone di disponibilità diverse.
18. È possibile selezionare uno o più gruppi di sicurezza a cui consentire l'accesso al cluster (ad esempio, i gruppi di sicurezza delle macchine client). Se specifichi gruppi di sicurezza condivisi con te, devi assicurarti di disporre delle autorizzazioni per utilizzarli. Nello specifico, è necessaria l'autorizzazione `ec2:DescribeSecurityGroups`. Connessione a un cluster Amazon MSK.
19. Seleziona Avanti.
20. Seleziona i metodi di controllo degli accessi e le impostazioni di crittografia del cluster per crittografare i dati durante il transito tra clienti e broker. Per ulteriori informazioni, consulta the section called “Crittografia in transito”.
21. Scegli il tipo di chiave KMS che desideri utilizzare per la crittografia dei dati a riposo. Per ulteriori informazioni, consulta the section called “Crittografia a riposo”.
22. Seleziona Successivo.
23. Scegli il monitoraggio e i tag che desideri. Questo determina il set di parametri che si ottiene. Per ulteriori informazioni, consulta Monitoraggio di un cluster. Amazon CloudWatch, Prometheus, Broker log delivery o Cluster tags, quindi seleziona Avanti.

24. Controlla le impostazioni del tuo cluster. Puoi tornare indietro e modificare le impostazioni selezionando Precedente per tornare alla schermata precedente della console o Modifica per modificare impostazioni specifiche del cluster. Se le impostazioni sono corrette, seleziona Crea cluster.
25. Verifica lo Stato del cluster nella pagina Riepilogo del cluster. Quando Amazon MSK assegna il cluster, lo stato passa da Creazione in corso ad Attivo. Quando lo stato è Attivo, puoi connetterti al cluster. Per ulteriori informazioni sugli stati del cluster, consulta la pagina Stati dei cluster.

Creazione di un cluster utilizzando AWS CLI

1. Copiare il JSON seguente e salvarlo in un file. Assegnare un nome al file `brokernodegroupinfo.json`. Sostituire gli ID di sottorete nel JSON con i valori corrispondenti alle sottoreti. Le sottoreti devono trovarsi in zone di disponibilità differenti. Sostituire *"Security-Group-ID"* con l'ID di uno o più gruppi di sicurezza del client VPC. I client associati a questi gruppi di sicurezza ottengono l'accesso al cluster. Se specifichi gruppi di sicurezza condivisi con te, devi verificare di disporre delle autorizzazioni per gli stessi. Nello specifico, è necessaria l'autorizzazione `ec2:DescribeSecurityGroups`. Per un esempio, consulta la pagina Amazon EC2: Allows Managing Amazon EC2 Security Groups Associated With a Specific VPC, Programmatically and in the Console. Infine, salva il file JSON aggiornato sul computer in cui è AWS CLI installato.

```
{
  "InstanceType": "kafka.m5.large",
  "ClientSubnets": [
    "Subnet-1-ID",
    "Subnet-2-ID"
  ],
  "SecurityGroups": [
    "Security-Group-ID"
  ]
}
```

> **Important**
>
> Se si utilizza la Regione Stati Uniti occidentali (California settentrionale), specificare esattamente due sottoreti. Per altre Regioni in cui Amazon MSK è disponibile, è possibile specificare due o tre sottoreti. Le sottoreti specificate devono trovarsi in zone

di disponibilità distinte. Quando crei un cluster, Amazon MSK distribuisce i nodi broker in modo uniforme sulle sottoreti specificate.

2. Esegui il AWS CLI comando seguente nella directory in cui hai salvato il `brokernodegroupinfo.json` file, sostituendo *«Your-Cluster-Name» con un nome* a tua scelta. Per *"Monitoring-Level"*, è possibile specificare uno dei seguenti tre valori: `DEFAULT`, `PER_BROKER` o `PER_TOPIC_PER_BROKER`. Per informazioni su questi tre diversi livelli di monitoraggio, consulta ???. Il parametro `enhanced-monitoring` è facoltativo. Se non viene specificato nel comando `create-cluster`, si ottiene il livello di monitoraggio `DEFAULT`.

```
aws kafka create-cluster --cluster-name "Your-Cluster-Name" --broker-node-group-
info file://brokernodegroupinfo.json --kafka-version "2.8.1" --number-of-broker-
nodes 3 --enhanced-monitoring "Monitoring-Level"
```

L'output del comando è simile al JSON seguente:

```
{
    "ClusterArn": "...",
    "ClusterName": "AWSKafkaTutorialCluster",
    "State": "CREATING"
}
```

> **Note**
>
> Il comando `create-cluster` potrebbe restituire un errore che indica che una o più sottoreti appartengono a zone di disponibilità non supportate. Quando ciò si verifica, l'errore indica quali zone di disponibilità non sono supportate. Crea sottoreti che non utilizzano le zone di disponibilità non supportate e riprova a eseguire nuovamente il comando `create-cluster`.

3. Salvare il valore della chiave `ClusterArn` perché è necessario per eseguire altre operazioni nel cluster.
4. Eseguire il comando seguente per verificare il tuo cluster `STATE`. Il valore `STATE` cambia da `CREATING` a `ACTIVE` quando Amazon EMR assegna il cluster. Quando lo stato è `ACTIVE`, puoi connetterti al cluster. Per ulteriori informazioni sugli stati del cluster, consulta la pagina Stati dei cluster.

```
aws kafka describe-cluster --cluster-arn <your-cluster-ARN>
```

Creazione di un cluster con una configurazione Amazon MSK personalizzata utilizzando AWS CLI

Per informazioni sulle configurazioni personalizzate di Amazon MSK e su come crearle, consulta la sezione Configurazione.

1. Salva il JSON seguente in un file, sostituendo *configuration-arn* con l'ARN della configurazione che desideri utilizzare per creare il cluster.

```
{
    "Arn": configuration-arn,
    "Revision": 1
}
```

2. Esegui il comando `create-cluster` e utilizza l'opzione `configuration-info` per puntare al file JSON salvato nella fase precedente. Di seguito è riportato un esempio.

```
aws kafka create-cluster --cluster-name ExampleClusterName --broker-node-group-
info file://brokernodegroupinfo.json --kafka-version "2.8.1" --number-of-broker-
nodes 3 --enhanced-monitoring PER_TOPIC_PER_BROKER --configuration-info file://
configuration.json
```

Di seguito è riportato un esempio di una risposta corretta dopo l'esecuzione di questo comando.

```
{
    "ClusterArn": "arn:aws:kafka:us-east-1:123456789012:cluster/
CustomConfigExampleCluster/abcd1234-abcd-dcba-4321-a1b2abcd9f9f-2",
    "ClusterName": "CustomConfigExampleCluster",
    "State": "CREATING"
}
```

Creazione di un cluster tramite l'API

Per creare un cluster utilizzando l'API, consulta CreateCluster.

Eliminazione di un cluster Amazon MSK

> **Note**
>
> Se il tuo cluster ha una policy di dimensionamento automatico, ti consigliamo di rimuovere la policy prima di eliminare il cluster. Per ulteriori informazioni, consulta Scalabilità automatica.

Eliminazione di un cluster utilizzando AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli il cluster MSK da eliminare selezionando la casella di controllo accanto ad esso.
3. Scegli Elimina e conferma l'eliminazione.

Eliminazione di un cluster utilizzando AWS CLI

Esegui il comando seguente, sostituendolo *ClusterArn* con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

```
aws kafka delete-cluster --cluster-arn ClusterArn
```

Eliminazione di un cluster tramite l'API

Per eliminare un cluster utilizzando l'API, consulta DeleteCluster.

Recupero dei broker di bootstrap per un cluster Amazon MSK

Ottenere i broker bootstrap utilizzando il AWS Management Console

Il termine broker di bootstrap si riferisce a un elenco di broker che un client Apache Kafka può utilizzare come punto di partenza per connettersi al cluster. Questo elenco non include necessariamente tutti i broker di un cluster.

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.

2. La tabella mostra tutti i cluster per la regione corrente in questo account. Scegli il nome di un cluster per visualizzarne la descrizione.
3. Nella pagina Riepilogo del cluster, scegli Visualizza informazioni sul client. Questo mostra i broker bootstrap e la stringa di connessione Apache. ZooKeeper

Ottenere i broker bootstrap utilizzando il AWS CLI

Esegui il comando seguente, sostituendolo *ClusterArn*con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

```
aws kafka get-bootstrap-brokers --cluster-arn ClusterArn
```

Per un cluster MSK che utilizza the section called “Controllo degli accessi IAM”, l'output di questo comando è simile all'esempio JSON seguente.

```
{
    "BootstrapBrokerStringSaslIam": "b-1.myTestCluster.123z8u.c2.kafka.us-
west-1.amazonaws.com:9098,b-2.myTestCluster.123z8u.c2.kafka.us-
west-1.amazonaws.com:9098"
}
```

L'esempio seguente mostra i broker di bootstrap per un cluster con accesso pubblico attivato. Usa il `BootstrapBrokerStringPublicSaslIam` per l'accesso pubblico e la `BootstrapBrokerStringSaslIam` stringa per l'accesso dall'interno AWS.

```
{
    "BootstrapBrokerStringPublicSaslIam": "b-2-public.myTestCluster.v4ni96.c2.kafka-
beta.us-east-1.amazonaws.com:9198,b-1-public.myTestCluster.v4ni96.c2.kafka-
beta.us-east-1.amazonaws.com:9198,b-3-public.myTestCluster.v4ni96.c2.kafka-beta.us-
east-1.amazonaws.com:9198",
    "BootstrapBrokerStringSaslIam": "b-2.myTestCluster.v4ni96.c2.kafka-
beta.us-east-1.amazonaws.com:9098,b-1.myTestCluster.v4ni96.c2.kafka-beta.us-
east-1.amazonaws.com:9098,b-3.myTestCluster.v4ni96.c2.kafka-beta.us-
east-1.amazonaws.com:9098"
}
```

La stringa dei broker di bootstrap deve contenere tre broker provenienti da tutte le zone di disponibilità in cui è implementato il cluster MSK (a meno che non siano disponibili solo due broker).

Recupero dei broker di bootstrap tramite l'API

Per far sì che i broker bootstrap utilizzino l'API, vedi GetBootstrap Brokers.

Elencazione dei cluster Amazon MSK

Elencare i cluster utilizzando il AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. La tabella mostra tutti i cluster per la regione corrente in questo account. Scegli il nome di un cluster per visualizzarne i dettagli.

Elenco dei cluster utilizzando il AWS CLI

Esegui il comando seguente.

```
aws kafka list-clusters
```

Elencazione dei cluster tramite l'API

Per elencare i cluster che utilizzano l'API, consulta. ListClusters

Gestione dei metadati

Amazon MSK supporta le modalità di gestione dei metadati Apache ZooKeeper o KRAFT.

Dalla versione 3.7.x di Apache Kafka su Amazon MSK, puoi creare cluster che utilizzano la modalità KRAFT anziché la modalità. ZooKeeper I cluster basati su Kraft si basano su controller all'interno di Kafka per gestire i metadati.

Argomenti

- ZooKeeper modalità
- modalità KRAFT

ZooKeeper modalità

Apache ZooKeeper è «un servizio centralizzato per la gestione delle informazioni di configurazione, la denominazione, la sincronizzazione distribuita e la fornitura di servizi di gruppo. Tutti questi tipi di servizi vengono utilizzati in una forma o nell'altra da applicazioni distribuite», incluso Apache Kafka.

Se il tuo cluster utilizza la ZooKeeper modalità, puoi utilizzare i passaggi seguenti per ottenere la stringa di connessione ZooKeeper Apache. Tuttavia, ti consigliamo di utilizzare il `BootstrapServerString` per connetterti al tuo cluster ed eseguire operazioni di amministrazione poiché il `--zookeeper` flag è stato reso obsoleto in Kafka 2.5 ed è stato rimosso da Kafka 3.0.

ZooKeeper Ottenere la stringa di connessione di Apache utilizzando il AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. La tabella mostra tutti i cluster per la regione corrente in questo account. Scegli il nome di un cluster per visualizzarne la descrizione.
3. Nella pagina Riepilogo del cluster, scegli Visualizza informazioni sul client. Questo mostra i broker bootstrap e la stringa di connessione ZooKeeper Apache.

Ottenere la stringa di connessione Apache usando ZooKeeper il AWS CLI

1. Se l'Amazon Resource Name (ARN) del cluster non è noto, puoi trovarlo elencando tutti i cluster nell'account. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.
2. Per ottenere la stringa di ZooKeeper connessione Apache, insieme ad altre informazioni sul cluster, esegui il comando seguente, sostituendolo *`ClusterArn`* con l'ARN del cluster.

```
aws kafka describe-cluster --cluster-arn ClusterArn
```

L'output di questo comando `describe-cluster` è simile all'esempio JSON seguente.

```
{
    "ClusterInfo": {
        "BrokerNodeGroupInfo": {
            "BrokerAZDistribution": "DEFAULT",
            "ClientSubnets": [
                "subnet-0123456789abcdef0",
                "subnet-2468013579abcdef1",
```

```
                "subnet-1357902468abcdef2"
            ],
            "InstanceType": "kafka.m5.large",
            "StorageInfo": {
                "EbsStorageInfo": {
                    "VolumeSize": 1000
                }
            }
        },
        "ClusterArn": "arn:aws:kafka:us-east-1:111122223333:cluster/
testcluster/12345678-abcd-4567-2345-abcdef123456-2",
        "ClusterName": "testcluster",
        "CreationTime": "2018-12-02T17:38:36.75Z",
        "CurrentBrokerSoftwareInfo": {
            "KafkaVersion": "2.2.1"
        },
        "CurrentVersion": "K13V1IB3VIYZZH",
        "EncryptionInfo": {
            "EncryptionAtRest": {
                "DataVolumeKMSKeyId": "arn:aws:kms:us-
east-1:555555555555:key/12345678-abcd-2345-ef01-abcdef123456"
            }
        },
        "EnhancedMonitoring": "DEFAULT",
        "NumberOfBrokerNodes": 3,
        "State": "ACTIVE",
        "ZookeeperConnectString": "10.0.1.101:2018,10.0.2.101:2018,10.0.3.101:2018"
    }
}
```

L'esempio JSON precedente mostra la chiave `ZookeeperConnectString` nell'output del comando `describe-cluster`. Copia il valore corrispondente a questa chiave e salvalo per utilizzarlo quando è necessario creare un argomento nel cluster.

> **Important**
>
> Il cluster Amazon MSK deve trovarsi nello `ACTIVE` stato in cui è possibile ottenere la stringa di ZooKeeper connessione Apache. Quando un cluster è ancora nello stato `CREATING`, l'output del comando `describe-cluster` non include `ZookeeperConnectString`. In questo caso, occorre attendere alcuni minuti ed

eseguire nuovamente `describe-cluster` dopo che il cluster raggiunge lo stato `ACTIVE`.

Ottenere la stringa di ZooKeeper connessione Apache tramite l'API

Per ottenere la stringa di ZooKeeper connessione Apache utilizzando l'API, vedi. DescribeCluster

modalità KRAFT

Amazon MSK ha introdotto il supporto per KRAFT (Apache Kafka Raft) nella versione 3.7.x di Kafka. La community di Apache Kafka ha sviluppato KRAFT per sostituire Apache per la gestione dei metadati nei cluster Apache Kafka. ZooKeeper In modalità KRAFT, i metadati del cluster vengono propagati all'interno di un gruppo di controller Kafka, che fanno parte del cluster Kafka, anziché tra i nodi. ZooKeeper I controller Kraft sono inclusi senza costi aggiuntivi per l'utente e non richiedono alcuna configurazione o gestione aggiuntiva da parte dell'utente. Vedi KIP-500 per ulteriori informazioni su KRAft.

Ecco alcuni punti da tenere in considerazione sulla modalità KRAFT su MSK:

- La modalità KRAFT è disponibile solo per i nuovi cluster. Non è possibile cambiare modalità di metadati una volta creato il cluster.
- Sulla console MSK, è possibile creare un cluster basato su Kraft scegliendo la versione 3.7.x di Kafka e selezionando la casella di controllo KRAFT nella finestra di creazione del cluster.
- Per creare un cluster in modalità Kraft utilizzando l'API o le operazioni MSK, è necessario utilizzare come versione. `CreateClusterCreateClusterV2``3.7.x.kraft` Usa `3.7.x` come versione per creare un cluster in ZooKeeper modalità.
- Il numero di partizioni per broker è lo stesso su Kraft e sui cluster ZooKeeper basati. Tuttavia, KRAFT consente di ospitare più partizioni per cluster fornendo più broker in un cluster.
- Non sono necessarie modifiche all'API per utilizzare la modalità KRAFT su Amazon MSK. Tuttavia, se i tuoi client utilizzano ancora la stringa di `--zookeeper` connessione oggi, dovresti aggiornarli in modo che utilizzino la stringa di `--bootstrap-server` connessione per connettersi al cluster. Il `--zookeeper` flag è obsoleto nella versione 2.5 di Apache Kafka e viene rimosso a partire dalla versione 3.0 di Kafka. Ti consigliamo quindi di utilizzare le versioni recenti del client Apache Kafka e la stringa di connessione per tutte le connessioni al tuo cluster. `--bootstrap-server`

- ZooKeeper la modalità continua a essere disponibile per tutte le versioni rilasciate in cui zookeeper è supportato anche da Apache Kafka. Vedi Versioni di Apache Kafka supportate i dettagli sulla fine del supporto per le versioni di Apache Kafka e gli aggiornamenti futuri.
- È necessario verificare che tutti gli strumenti utilizzati siano in grado di utilizzare le API di amministrazione di Kafka senza connessioni. ZooKeeper Consulta la procedura aggiornata Utilizzo LinkedIn del Cruise Control per Apache Kafka con Amazon MSK per connettere il cluster a Cruise Control. Cruise Control fornisce anche istruzioni per utilizzare il Cruise Control senza ZooKeeper.
- Non è necessario accedere direttamente ai controller KRAFT del cluster per eventuali azioni amministrative. Tuttavia, se utilizzate il monitoraggio aperto per raccogliere le metriche, avete bisogno anche degli endpoint DNS dei vostri controller per raccogliere alcune metriche non relative ai controller sul vostro cluster. È possibile ottenere questi endpoint DNS dalla console MSK o utilizzando l'operazione API. ListNodes Vedi i passaggi aggiornati Monitoraggio aperto con Prometheus per configurare il monitoraggio aperto per i cluster basati su Kraft.
- Non ci sono CloudWatch metriche aggiuntive da monitorare per i cluster in modalità Kraft rispetto ai cluster modali. ZooKeeper MSK gestisce i controller KRAFT utilizzati nei cluster.
- È possibile continuare a gestire gli ACL utilizzando i cluster in modalità Kraft utilizzando la stringa di connessione. `--bootstrap-server` Non è necessario utilizzare la stringa di `--zookeeper` connessione per gestire gli ACL. Per informazioni, consulta ACL Apache Kafka.
- In modalità Kraft, i metadati del cluster vengono archiviati sui controller KRAFT all'interno di Kafka e non su nodi esterni. ZooKeeper Pertanto, non è necessario controllare l'accesso ai nodi del controller separatamente come si fa con i nodi. ZooKeeper

Gestione dello storage

Amazon MSK offre funzionalità per aiutarti con la gestione dell'archiviazione sui tuoi cluster MSK.

Argomenti

- Archiviazione a più livelli
- Aumento delle dimensioni dello spazio di archiviazione del broker
- Assegnazione della velocità di trasmissione effettiva dell'archiviazione

Archiviazione a più livelli

L'archiviazione a più livelli è un livello di archiviazione a basso costo per Amazon MSK che si dimensiona fino a una capacità praticamente illimitata, rendendo conveniente la creazione di applicazioni di streaming di dati.

È possibile creare un cluster Amazon MSK configurato con un'archiviazione a più livelli che bilancia prestazioni e costi. Amazon MSK archivia i dati in streaming in un livello di archiviazione primario ottimizzato per le prestazioni fino a raggiungere i limiti di conservazione degli argomenti di Apache Kafka. Quindi, Amazon MSK sposta automaticamente i dati nel nuovo livello di archiviazione a basso costo.

Quando l'applicazione inizia a leggere i dati dall'archiviazione a più livelli, è possibile che i primi byte siano soggetti a un aumento della latenza di lettura. Quando inizi a leggere i dati rimanenti in sequenza dal livello a basso costo, le latenze dovrebbero essere simili a quelle del livello di archiviazione primario. Non è necessario effettuare il provisioning di alcun tipo di archiviazione per l'archiviazione più livelli a basso costo o per gestire l'infrastruttura. È possibile archiviare qualsiasi quantità di dati e pagare solo per le risorse utilizzate. Questa funzionalità è compatibile con le API introdotte in KIP-405: Kafka Tiered Storage.

Di seguito sono elencate alcune caratteristiche dell'archiviazione a più livelli:

- È possibile dimensionare fino a una capacità di archiviazione praticamente illimitata. Non è necessario fare supposizioni su come dimensionare la propria infrastruttura Apache Kafka.
- È possibile mantenere i dati più a lungo negli argomenti di Apache Kafka o aumentare lo spazio di archiviazione degli argomenti senza la necessità di aumentare il numero di broker.
- Fornisce un buffer di sicurezza di maggiore durata per gestire ritardi imprevisti nell'elaborazione.
- Puoi rielaborare i vecchi dati nel loro esatto ordine di produzione con il codice di elaborazione del flusso esistente e le API di Kafka.
- Le partizioni si ribilanciano più velocemente perché i dati nell'archiviazione secondaria non richiedono la replica tra i dischi del broker.
- I dati tra i broker e l'archiviazione a più livelli si spostano all'interno del VPC e non viaggiano su Internet.
- Per connettersi a nuovi cluster con l'archiviazione a più livelli abilitata, un computer client può utilizzare lo stesso processo che utilizza per connettersi a un cluster senza l'archiviazione a più livelli abilitata. Consulta la sezione Creazione di un computer client.

Requisiti di archiviazione a più livelli

- È necessario utilizzare la versione 3.0.0 o successiva del client Apache Kafka per creare un nuovo argomento con l'archiviazione a più livelli abilitata. Per trasferire un argomento esistente all'archiviazione a più livelli, puoi riconfigurare un computer client che utilizza una versione del client Kafka precedente alla 3.0.0 (la versione minima supportata di Apache Kafka è 2.8.2.tiered) per abilitare l'archiviazione a più livelli. Per informazioni, consulta Passaggio 4: creazione di un argomento.
- Il cluster Amazon MSK con storage su più livelli abilitato deve utilizzare la versione 3.6.0 o successiva o 2.8.2.tiered.

Vincoli e limitazioni dell'archiviazione a più livelli

L'archiviazione a più livelli presenta i seguenti vincoli e limitazioni:

- L'archiviazione a più livelli si applica solo ai cluster in modalità assegnata.
- Lo storage su più livelli non supporta la dimensione del broker t3.small.
- Il periodo di conservazione minimo nell'archiviazione a basso costo è di 3 giorni. Non è previsto un periodo minimo di conservazione per l'archiviazione primaria.
- L'archiviazione a più livelli non supporta le directory di log multipli su un broker (funzionalità relative a JBOD).
- L'archiviazione a più livelli non supporta gli argomenti compressi. Assicurati che cleanup.policy sia configurata solo su "DELETE" per tutti gli argomenti per cui è attivata l'archiviazione a più livelli.
- L'archiviazione a più livelli può essere disabilitata per singoli argomenti ma non per l'intero cluster. Una volta disattivata, l'archiviazione a più livelli non può essere riattivata per un argomento.
- Se utilizzi la versione 2.8.2.tiered di Amazon MSK, puoi migrare solo a un'altra versione di Apache Kafka supportata dallo storage su più livelli. Se non desideri continuare a utilizzare una versione supportata dallo storage su più livelli, crea un nuovo cluster MSK e migra i tuoi dati su di esso.
- Lo kafka-log-dirs strumento non è in grado di riportare le dimensioni dei dati di storage su più livelli. Lo strumento riporta solo la dimensione dei segmenti di log nell'archiviazione primaria.

Come vengono copiati i segmenti di log nell'archiviazione a più livelli

Quando abiliti l'archiviazione a più livelli per un argomento nuovo o esistente, Apache Kafka copia i segmenti di log chiusi dall'archiviazione primaria all'archiviazione a più livelli.

- Apache Kafka copia solo i segmenti di log chiusi. Copia tutti i messaggi all'interno del segmento di log in un'archiviazione a più livelli.
- I segmenti attivi non sono idonei per l'archiviazione a più livelli. La dimensione del segmento di log (segment.bytes) o il tempo di distribuzione del segmento (segment.ms) controllano la velocità di chiusura dei segmenti e la velocità con cui, successivamente, Apache Kafka li copia nell'archiviazione a più livelli.

Le impostazioni di conservazione per un argomento con l'archiviazione a più livelli abilitata sono diverse dalle impostazioni per un argomento senza l'archiviazione a più livelli abilitata. Le seguenti regole disciplinano la conservazione dei messaggi negli argomenti con l'archiviazione a più livelli abilitata:

- È possibile definire la conservazione in Apache Kafka con due impostazioni: log.retention.ms (durata) e log.retention.bytes (dimensioni). Queste impostazioni determinano la durata e le dimensioni totali dei dati che Apache Kafka conserva nel cluster. Indipendentemente dal fatto che si abiliti o meno la modalità di archiviazione a più livelli, queste configurazioni vengono impostate a livello di cluster. È possibile sovrascrivere le impostazioni a livello di argomento con le configurazioni degli argomenti.
- Quando si abilita l'archiviazione a più livelli, è possibile specificare anche per quanto tempo il livello di archiviazione primaria ad alte prestazioni archivia i dati. Ad esempio, se un argomento ha un'impostazione di conservazione complessiva (log.retention.ms) di 7 giorni e una conservazione locale (local.retention.ms) di 12 ore, l'archiviazione primaria del cluster conserva i dati solo per le prime 12 ore. Il livello di archiviazione a basso costo conserva i dati per tutti i 7 giorni.
- Al log completo si applicano le normali impostazioni di conservazione. Ciò include le parti primarie e a più livelli.
- Le impostazioni local.retention.ms o local.retention.bytes controllano la conservazione dei messaggi nell'archiviazione primaria. Quando i dati hanno raggiunto le soglie di impostazione della conservazione dell'archiviazione primaria (local.retention.ms/bytes) su un log completo, Apache Kafka copia i dati nell'archiviazione primaria a più livelli. I dati sono quindi idonei alla scadenza.
- Quando Apache Kafka copia un messaggio in un segmento di log a più livelli, lo rimuove dal cluster in base alle impostazioni retention.ms o retention.bytes.

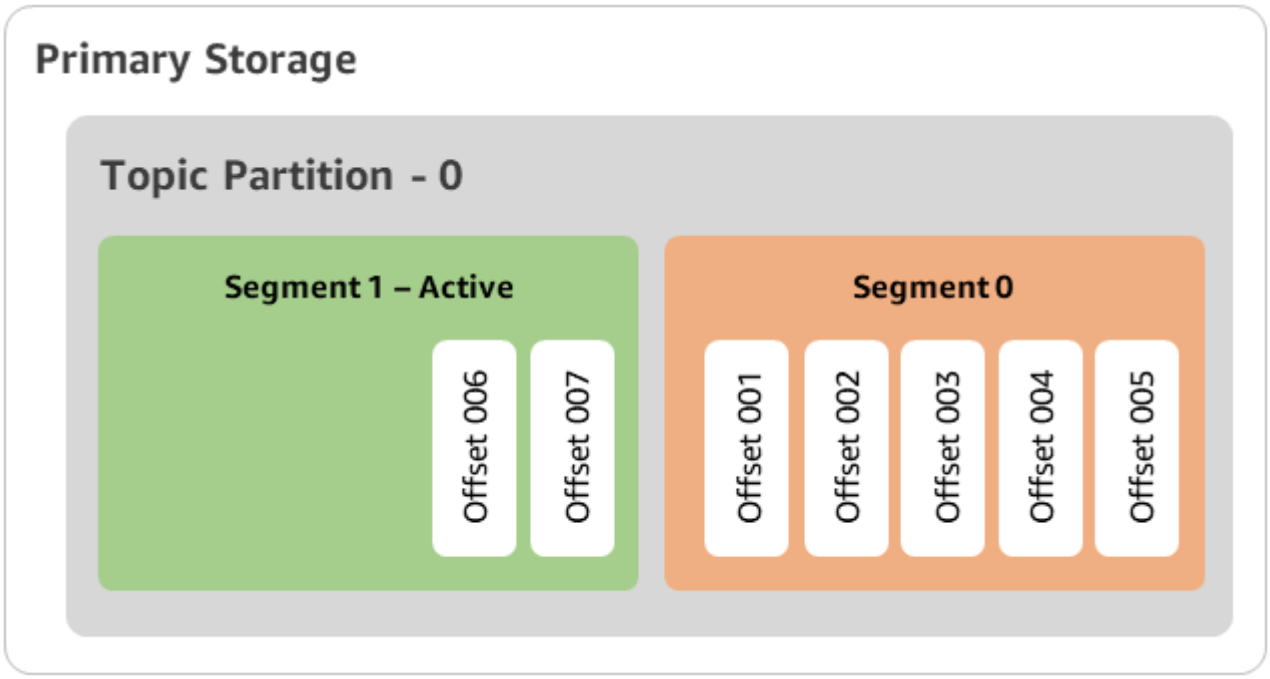
Esempio di scenario di archiviazione a più livelli

Questo scenario illustra il comportamento di un argomento esistente che contiene messaggi nell'archiviazione primaria quando è abilitata l'archiviazione a più livelli. L'archiviazione a più livelli

su questo argomento viene abilitata quando si imposta remote.storage.enable su `true`. In questo esempio, retention.ms è impostato su 5 giorni e local.retention.ms è impostato su 2 giorni. Di seguito è riportata la sequenza di eventi alla scadenza di un segmento.

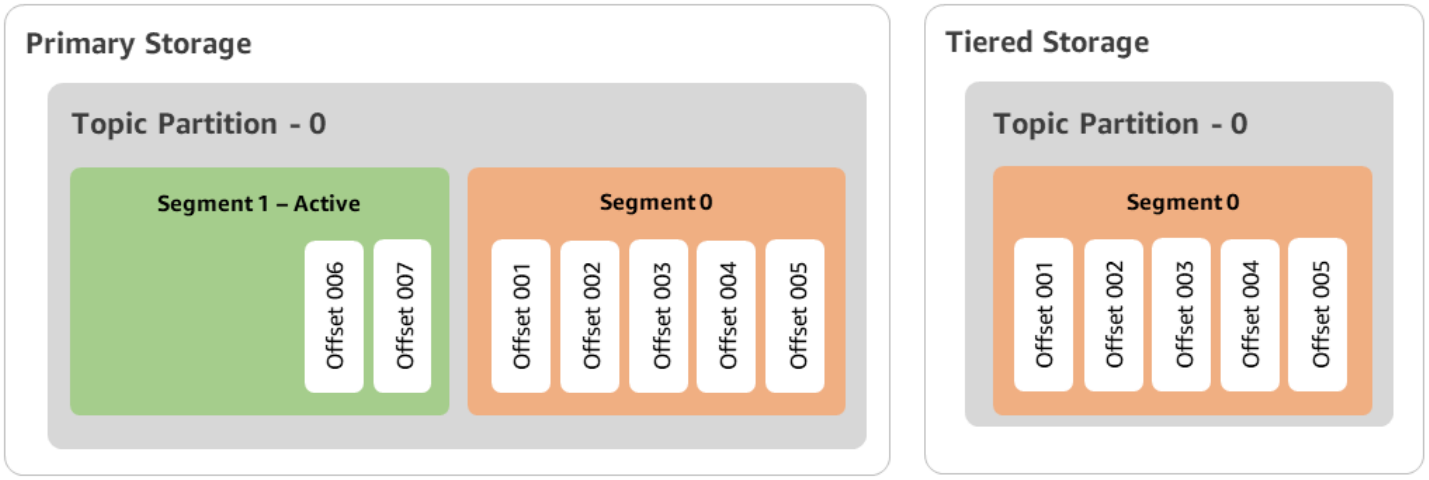
Ora T0: prima di abilitare l'archiviazione a più livelli.

Prima di abilitare l'archiviazione a più livelli per questo argomento, esistono due segmenti di log. Uno dei segmenti è attivo per una partizione di argomenti esistente 0.



Ora T1 (< 2 giorni): archiviazione a più livelli abilitata. Segmento 0 copiato nell'archiviazione a più livelli.

Dopo aver abilitato l'archiviazione a più livelli per questo argomento, Apache Kafka copia il segmento di log 0 nell'archiviazione a più livelli dopo che il segmento soddisfa le impostazioni di conservazione iniziali. Apache Kafka conserva anche la copia di archiviazione primaria del segmento 0. Il segmento 1 attivo non è ancora idoneo alla copia nell'archiviazione a più livelli. In questa sequenza temporale, Amazon MSK non applica ancora nessuna delle impostazioni di conservazione per alcuno dei messaggi nel segmento 0 e nel segmento 1 (local.retention.bytes/ms, retention.ms/bytes)



Ora T2: conservazione locale in vigore.

Dopo 2 giorni, le impostazioni di conservazione dell'archiviazione primaria hanno effetto per il segmento 0 che Apache Kafka ha copiato nell'archiviazione a più livelli. Ciò è determinato dall'impostazione di local.retention.ms su 2 giorni. Il segmento 0 ora scade dall'archiviazione primaria. Il segmento 1 è attivo, pertanto non è ancora idoneo né alla scadenza né a essere copiato nell'archiviazione a più livelli.

Ora T3: conservazione complessiva in vigore.

Dopo 5 giorni, le impostazioni di conservazione hanno effetto e Kafka cancella il segmento di log 0 e i messaggi associati dall'archiviazione a più livelli. Il segmento 1 non è ancora idoneo alla scadenza né può essere copiato nell'archiviazione a più livelli perché è attivo. Il segmento 1 non è ancora chiuso, quindi non è idoneo per la distribuzione dei segmenti.

Creazione di un cluster Amazon MSK con storage su più livelli con AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli Create cluster (Crea cluster).
3. Scegli Creazione personalizzata per l'archiviazione a più livelli.
4. Specificare un nome per il cluster.
5. In Tipo di cluster, seleziona Assegnato.
6. Scegli la versione di Amazon Kafka che supporti l'archiviazione a più livelli e che desideri che Amazon MSK utilizzi per creare il cluster.
7. Specificate una dimensione del broker diversa da kafka.t3.small.
8. Specifica il numero di broker che devono essere creati da Amazon MSK in ogni zona di disponibilità. Il valore minimo è un broker per zona di disponibilità e il valore massimo è 30 broker per cluster.
9. Specifica il numero di zone in cui sono distribuiti i broker.
10. Specifica il numero di broker Apache Kafka implementati per zona.
11. Seleziona Opzioni di archiviazione. Ciò include l'archiviazione a più livelli e l'archiviazione EBS per abilitare la modalità di archiviazione a più livelli.
12. Segui i restanti passaggi nella procedura guidata di creazione dei cluster. Al termine, Archiviazione a più livelli e archiviazione EBS viene visualizzata come modalità di archiviazione del cluster nella vista Rivedi e crea.
13. Selezionare Creazione di un cluster.

Creazione di un cluster Amazon MSK con storage su più livelli con AWS CLI

Per abilitare l'archiviazione a più livelli su un cluster, crea il cluster con la versione e l'attributo di Apache Kafka corretti per l'archiviazione a più livelli. Segui l'esempio di codice sottostante. Inoltre, completa la procedura descritta nella sezione successiva per Creazione di un argomento su Kafka con l'archiviazione a più livelli abilitata.

Per un elenco completo degli attributi supportati per la creazione di cluster, consulta la sezione create-cluster.

```
aws tiered-storage create-cluster \
 –cluster-name "MessagingCluster" \
 –broker-node-group-info file://brokernodegroupinfo.json \
 –number-of-broker-nodes 3 \
--kafka-version "3.6.0" \
--storage-mode "TIERED"
```

Creazione di un argomento su Kafka con l'archiviazione a più livelli abilitata

Per completare il processo avviato quando hai creato un cluster con l'archiviazione a più livelli abilitata, crea anche un argomento con l'archiviazione a più livelli abilitata con gli attributi dell'esempio di codice successivo. Gli attributi specifici per l'archiviazione a più livelli sono i seguenti:

- `local.retention.ms` (ad esempio, 10 minuti) per le impostazioni di conservazione basate sul tempo o `local.retention.bytes` per i limiti delle dimensioni dei segmenti di log.
- `remote.storage.enable` impostato su `true` per abilitare l'archiviazione a più livelli.

La configurazione seguente utilizza local.retention.ms, ma è possibile sostituire questo attributo con local.retention.bytes. Questo attributo controlla la quantità di tempo che può trascorrere o il numero di byte che Apache Kafka può copiare prima che il servizio copi i dati dall'archiviazione primaria a quella a più livelli. Per maggiori dettagli sugli attributi di configurazione supportati, consulta la sezione Configurazione a livello di argomento.

> **Note**
>
> È necessario utilizzare la versione 3.0.0 o successiva del client Apache Kafka. Queste versioni supportano un'impostazione chiamata `remote.storage.enable` solo in tali versioni client di `kafka-topics.sh`. Per abilitare l'archiviazione a più livelli su un argomento

esistente che utilizza una versione precedente di Apache Kafka, consulta la sezione Abilitazione dell'archiviazione a più livelli su un argomento esistente.

```
bin/kafka-topics.sh --create --bootstrap-server $bs --replication-factor 2
 --partitions 6 --topic MSKTutorialTopic --config remote.storage.enable=true
 --config local.retention.ms=100000 --config retention.ms=604800000 --config
 segment.bytes=134217728
```

Abilitazione e disabilitazione dell'archiviazione a più livelli su un argomento esistente

Queste sezioni spiegano come abilitare e disabilitare l'archiviazione a più livelli su un argomento che hai già creato. Per creare un nuovo cluster e un argomento con l'archiviazione a più livelli abilitata, consulta la sezione Creazione di un cluster con archiviazione a più livelli tramite la AWS Management Console.

Abilitazione dell'archiviazione a più livelli su un argomento esistente

Per abilitare l'archiviazione a più livelli su un argomento esistente, utilizza la sintassi del comando `alter` nell'esempio seguente. Quando abiliti l'archiviazione a più livelli su un argomento esistente, non è necessario utilizzare una determinata versione del client Apache Kafka.

```
bin/kafka-configs.sh --bootstrap-server $bsrv --alter --entity-type topics
 --entity-name msk-ts-topic --add-config 'remote.storage.enable=true,
 local.retention.ms=604800000, retention.ms=15550000000'
```

Disabilitazione dell'archiviazione a più livelli su un argomento esistente

Per disabilitare l'archiviazione a più livelli su un argomento esistente, utilizza la sintassi del comando `alter` nello stesso ordine con cui hai abilitato l'archiviazione a più livelli.

```
bin/kafka-configs.sh --bootstrap-server $bs --alter --entity-type topics --
entity-name MSKTutorialTopic --add-config 'remote.log.msk.disable.policy=Delete,
 remote.storage.enable=false'
```

> **Note**
>
> Quando si disabilita l'archiviazione a più livelli, si eliminano completamente i dati relativi all'argomento nell'archiviazione a più livelli. Apache Kafka conserva i dati dell'archiviazione

primaria, ma applica comunque le regole di conservazione dell'archiviazione primaria in base a `local.retention.ms`. Una volta disabilitata l'archiviazione a più livelli su un argomento, non sarà possibile riabilitarla. Se desideri disabilitare l'archiviazione a più livelli su un argomento esistente, non è necessario utilizzare una determinata versione del client Apache Kafka.

Abilitazione dello storage su più livelli su un cluster esistente tramite AWS CLI

> **Note**
>
> È possibile abilitare l'archiviazione a più livelli solo se la policy del cluster log.cleanup.policy è impostata su `delete`, poiché gli argomenti compatti non sono supportati nell'archiviazione a più livelli. Successivamente, puoi configurare la policy log.cleanup.policy di un singolo argomento su `compact` in modo che l'archiviazione a più livelli non sia abilitata su quel particolare argomento. Per maggiori dettagli sugli attributi di configurazione supportati, consulta la sezione Configurazione a livello di argomento.

1. Aggiorna la versione di Kafka: le versioni dei cluster non sono semplici numeri interi. Per trovare la versione corrente del cluster, utilizzare l'`DescribeCluster`operazione o il comando `describe-cluster` AWS CLI. Una versione di esempio è `KTVPDKIKX0DER`.

```
aws kafka update-cluster-kafka-version --cluster-arn ClusterArn --current-version
 Current-Cluster-Version --target-kafka-version 3.6.0
```

2. Modifica la modalità di archiviazione del cluster. Nel seguente esempio di codice viene illustrato come modificare la modalità di archiviazione del cluster in `TIERED` tramite l'API `update-storage`.

```
aws kafka update-storage --current-version Current-Cluster-Version --cluster-arn
 Cluster-arn --storage-mode TIERED
```

Aggiornamento dell'archiviazione a più livelli su un cluster esistente tramite la console

> **Note**
>
> È possibile abilitare l'archiviazione a più livelli solo se la policy del cluster log.cleanup.policy è impostata su `delete`, poiché gli argomenti compatti non sono supportati nell'archiviazione a più livelli. Successivamente, puoi configurare la policy log.cleanup.policy di un singolo argomento su `compact` in modo che l'archiviazione a più livelli non sia abilitata su quel particolare argomento. Per maggiori dettagli sugli attributi di configurazione supportati, consulta la sezione Configurazione a livello di argomento.

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Vai alla pagina di riepilogo del cluster e scegli Proprietà.
3. Vai alla sezione Archiviazione e scegli Modifica modalità di archiviazione del cluster.
4. Scegli Archiviazione a più livelli e archiviazione EBS e Salva modifiche.

Aumento delle dimensioni dello spazio di archiviazione del broker

È possibile aumentare la quantità di storage EBS per broker. Non è possibile ridurre lo storage.

I volumi di storage rimangono disponibili durante questa operazione di dimensionamento.

> **Important**
>
> Quando l'archiviazione viene dimensionata per un cluster MSK, l'archiviazione aggiuntiva viene resa disponibile immediatamente. Tuttavia, il cluster richiede un periodo di raffreddamento dopo ogni evento di dimensionamento dell'archiviazione. Amazon MSK utilizza questo periodo di raffreddamento per ottimizzare il cluster prima di un successivo nuovo dimensionamento. Questo periodo può variare da un minimo di 6 ore a più di 24 ore, a seconda delle dimensioni e dell'utilizzo dell'archiviazione del cluster e del traffico. Ciò è applicabile sia agli eventi di ridimensionamento automatico che al ridimensionamento manuale utilizzando l'operazione di UpdateBrokerarchiviazione. Per informazioni sul corretto dimensionamento dell'archiviazione, consulta la sezione Best practice.

Puoi utilizzare l'archiviazione a più livelli per aumentare fino a quantità illimitate lo spazio di archiviazione per il broker. Per informazioni, consultare Archiviazione a più livelli.

Argomenti

- Scalabilità automatica
- Dimensionamento manuale

Scalabilità automatica

Per espandere automaticamente l'archiviazione del cluster in risposta a un maggiore utilizzo, puoi configurare una policy di dimensionamento automatico dell'applicazione per Amazon MSK. In una policy di dimensionamento automatico, si imposta l'utilizzo del disco di destinazione e la capacità di dimensionamento massima.

Prima di utilizzare il dimensionamento automatico per Amazon MSK, è consigliabile tenere in considerazione quanto segue:

- **Important**

  Un'operazione di dimensionamento dell'archiviazione può avvenire solo una volta ogni sei ore.

  Ti consigliamo di iniziare con un volume di archiviazione della dimensione giusta per le tue esigenze di archiviazione. Per indicazioni sul corretto dimensionamento del cluster, consulta la pagina Dimensionamento corretto del cluster: numero di broker per cluster.
- Amazon MSK non riduce lo spazio di archiviazione del cluster in risposta a un utilizzo ridotto. Amazon MSK non supporta la riduzione delle dimensioni dei volumi di archiviazione. Se è necessario ridurre le dimensioni dell'archiviazione del cluster, è necessario migrare il cluster esistente in un cluster con un'archiviazione più piccola. Per ulteriori informazioni sulla migrazione di un cluster, consulta la pagina Migrazione.
- Amazon MSK non supporta il dimensionamento automatico nelle regioni Asia Pacifico (Osaka-Locale) e Africa (Città del Capo).
- Quando associ una politica di auto-scaling al tuo cluster, Amazon EC2 Auto Scaling crea automaticamente un allarme Amazon per il tracciamento degli obiettivi. CloudWatch Se si elimina un cluster con una politica di auto-scaling, CloudWatch questo allarme persiste. Per eliminare l' CloudWatch allarme, è necessario rimuovere una politica di auto-scaling da un cluster

prima di eliminare il cluster. Per ulteriori informazioni sul monitoraggio degli obiettivi, consulta la pagina Target tracking scaling policies for Amazon EC2 Auto Scaling nella Guida per l'utente di Dimensionamento automatico Amazon EC2.

Dettagli della policy di dimensionamento automatico

Una policy di dimensionamento automatico definisce i seguenti parametri predefiniti per il cluster:

- Obiettivo di utilizzo dell'archiviazione: la soglia di utilizzo dell'archiviazione utilizzata da Amazon MSK per attivare un'operazione di dimensionamento automatico. È possibile impostare l'obiettivo di utilizzo tra il 10% e l'80% della capacità di archiviazione corrente. Consigliamo di impostare l'obiettivo di utilizzo dell'archiviazione tra il 50% e il 60%.
- Capacità massima di archiviazione: il limite di scalabilità massimo che Amazon MSK può impostare per l'archiviazione del broker. È possibile impostare la capacità di archiviazione massima fino a 16 TiB per broker. Per ulteriori informazioni, consulta Quota di Amazon MSK.

Quando Amazon MSK rileva che il parametro `Maximum Disk Utilization` è uguale o superiore all'impostazione `Storage Utilization Target`, aumenta la capacità di archiviazione di una quantità pari al più grande tra due numeri: 10 GiB o il 10% dell'archiviazione corrente. Ad esempio, se hai 1.000 GiB, tale quantità è 100 GiB. Il servizio verifica l'utilizzo dell'archiviazione ogni minuto. Ulteriori operazioni di dimensionamento continuano ad aumentare l'archiviazione di una quantità pari al più grande tra due numeri: 10 GiB o il 10% dell'archiviazione corrente.

Per determinare se sono state eseguite operazioni di auto-scaling, utilizzare l'operazione. ListClusterOperations

Configurazione del dimensionamento automatico per il cluster Amazon MSK

Puoi utilizzare la console Amazon MSK, l'API Amazon MSK o implementare il ridimensionamento automatico AWS CloudFormation per lo storage. CloudFormation il supporto è disponibile tramite. Application Auto Scaling

> **Note**
>
> Non è possibile implementare il dimensionamento automatico al momento della creazione di un cluster. È necessario innanzitutto creare il cluster, quindi creare e abilitare una policy di dimensionamento automatico per il cluster. Tuttavia, puoi creare la policy mentre il servizio Amazon MSK crea il tuo cluster.

Configurazione del dimensionamento automatico tramite la AWS Management Console

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Nell'elenco di cluster, scegli il tuo cluster. Questa operazione ti reindirizzerà a una pagina che elenca i dettagli sul cluster.
3. Nella sezione Dimensionamento automatico per l'archiviazione, scegli Configura.
4. Crea e assegna un nome a una policy di dimensionamento automatico. Specifica l'obiettivo di utilizzo dell'archiviazione, la capacità massima di archiviazione e il parametro obiettivo.
5. Scegli `Save changes`.

Quando salvi e abiliti la nuova policy, la policy diventa attiva per il cluster. Quando viene raggiunto l'obiettivo di utilizzo dell'archiviazione, Amazon MSK espande l'archiviazione del cluster.

Configurazione del dimensionamento automatico tramite la CLI

1. Utilizza il RegisterScalableTargetcomando per registrare un obiettivo di utilizzo dello storage.
2. Usa il PutScalingPolicycomando per creare una politica di espansione automatica.

Configurazione del dimensionamento automatico tramite l'API

1. Utilizza l' RegisterScalableTargetAPI per registrare un obiettivo di utilizzo dello storage.
2. Utilizza l' PutScalingPolicyAPI per creare una politica di espansione automatica.

Dimensionamento manuale

Per incrementare lo storage, attendere che lo stato del cluster diventi `ACTIVE`. Il dimensionamento dell'archiviazione prevede un periodo di raffreddamento di almeno sei ore tra un evento e l'altro. Anche se l'operazione rende immediatamente disponibile spazio di archiviazione aggiuntivo, il servizio esegue ottimizzazioni sul cluster che possono richiedere fino a 24 ore o più. La durata di queste ottimizzazioni è proporzionale alla dimensione dell'archiviazione.

Scalabilità dello storage del broker utilizzando il AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli il cluster MSK per cui desideri aggiornare lo spazio di archiviazione del broker.

3. Nella sezione Archiviazione, scegli Modifica.
4. Specifica il volume di storage desiderato. La quantità di storage può essere solo aumentata, non diminuita.
5. Seleziona Salvataggio delle modifiche.

Scalabilità dello storage dei broker utilizzando il AWS CLI

Esegui il comando seguente, sostituendolo *ClusterArn*con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

Sostituisci *Current-Cluster-Version* con la versione corrente del cluster.

> **Important**
>
> Le versioni del cluster non sono interi semplici. Per trovare la versione corrente del cluster, usa l'DescribeClusteroperazione o il comando AWS CLI describe-cluster. Una versione di esempio è `KTVPDKIKX0DER`.

Il parametro *Target-Volume-in-GiB* rappresenta la quantità di storage di cui deve disporre ogni broker. È consentito aggiornare lo storage solo per tutti i broker. Non è possibile specificare singoli broker per i quali aggiornare lo storage. Il valore specificato per *Target-Volume-in-GiB* deve essere un numero intero maggiore di 100 GiB. Lo storage per broker dopo l'operazione di aggiornamento non può superare 16384 GiB.

```
aws kafka update-broker-storage --cluster-arn ClusterArn --current-version Current-
Cluster-Version --target-broker-ebs-volume-info '{"KafkaBrokerNodeId": "All",
 "VolumeSizeGB": Target-Volume-in-GiB}'
```

Aumento delle dimensioni dello spazio di archiviazione del broker tramite l'API

Per aggiornare lo storage di un broker utilizzando l'API, vedi Storage. UpdateBroker

Assegnazione della velocità di trasmissione effettiva dell'archiviazione

I broker Amazon MSK mantengono i dati sui volumi di archiviazione. L'I/O dell'archiviazione viene utilizzato quando i produttori scrivono sul cluster, quando i dati vengono replicati tra broker e

quando i consumatori leggono dati che non sono in memoria. La velocità di trasmissione effettiva dell'archiviazione del volume è la velocità con cui i dati possono essere scritti e letti da un volume di archiviazione. La velocità di trasmissione effettiva dell'archiviazione assegnata è la capacità di specificare tale velocità per i broker del cluster.

È possibile specificare la velocità di throughput assegnata in MiB al secondo per i cluster i cui broker sono di dimensioni `kafka.m5.4xlarge` o superiori e se il volume di storage è pari o superiore a 10 GiB. È possibile specificare la velocità di trasmissione effettiva assegnata durante la creazione del cluster. Inoltre, è possibile abilitare o disabilitare la velocità di trasmissione effettiva assegnata per un cluster che si trova nello stato `ACTIVE`.

Colli di bottiglia nella velocità di trasmissione effettiva

I colli di bottiglia nella velocità di trasmissione effettiva dei broker sono dovuti a molteplici cause: velocità di trasmissione effettiva del volume, velocità di trasmissione effettiva della rete da Amazon EC2 ad Amazon EBS e velocità di trasmissione effettiva in uscita di Amazon EC2. È possibile abilitare la velocità di trasmissione effettiva assegnata per regolare la velocità di trasmissione effettiva del volume. Tuttavia, le limitazioni della velocità di trasmissione effettiva del broker possono essere causate dalla velocità di trasmissione effettiva della rete da Amazon EC2 ad Amazon EBS e dalla velocità di trasmissione effettiva in uscita di Amazon EC2.

La velocità di trasmissione effettiva in uscita di Amazon EC2 è influenzata dal numero di gruppi di consumatori e dal numero di consumatori per ciascun gruppo. Inoltre, sia il throughput di rete da Amazon EC2 ad Amazon EBS che il throughput di uscita di Amazon EC2 sono più elevati per broker di grandi dimensioni.

Per volumi di dimensioni pari o superiori a 10 GiB, è possibile assegnare una velocità di trasmissione effettiva dell'archiviazione pari o superiore a 250 MiB al secondo. L'impostazione predefinita è 250 MiB al secondo. Per effettuare il provisioning del throughput di storage, devi scegliere la dimensione del broker kafka.m5.4xlarge o superiore (oppure kafka.m7g.2xlarge o superiore) e puoi specificare il throughput massimo come mostrato nella tabella seguente.

dimensione del broker	Velocità di trasmissione effettiva massima (MiB/secondo)
kafka.m5.4xlarge	593
kafka.m5.8xlarge	850

dimensione del broker	Velocità di trasmissione effettiva massima (MiB/secondo)
kafka.m5.12xlarge	1000
kafka.m5.16xlarge	1000
kafka.m5.24xlarge	1000
kafka.m7 g. 2 x grande	312,5
kafka.m7g.4xlarge	625
kafka.m7g.8xlarge	1000
kafka.m7g. 12 x grande	1000
kafka.m7g. 16 x grande	1000

Misurazione della velocità di trasmissione effettiva dell'archiviazione

È possibile utilizzare i parametri `VolumeReadBytes` e `VolumeWriteBytes` per misurare la velocità di trasmissione effettiva media di archiviazione di un cluster. La somma di questi due parametri fornisce la velocità di trasmissione effettiva media dell'archiviazione espressa in byte. Per ottenere la velocità di trasmissione effettiva media dell'archiviazione per un cluster, imposta questi due parametri su SUM e il periodo su 1 minuto, quindi utilizza la formula seguente.

```
Average storage throughput in MiB/s = (Sum(VolumeReadBytes) + Sum(VolumeWriteBytes)) /
 (60 * 1024 * 1024)
```

Per ulteriori informazioni sui parametri `VolumeReadBytes` e `VolumeWriteBytes`, consulta la sezione the section called “Monitoraggio del livello PER_BROKER”.

Aggiornamento della configurazione

Puoi aggiornare la configurazione di Amazon MSK prima o dopo aver attivato la velocità di trasmissione effettiva assegnata. Tuttavia, non vedrai la velocità di trasmissione effettiva desiderata finché non eseguirai entrambe le operazioni: aggiornare il parametro di configurazione `num.replica.fetchers` e attivare la velocità di trasmissione effettiva assegnata.

Nella configurazione predefinita di Amazon MSK, `num.replica.fetchers` ha un valore di 2. Per aggiornare il `num.replica.fetchers`, puoi utilizzare i valori suggeriti dalla tabella seguente. Questi valori sono forniti a scopo indicativo. Si consiglia di modificare questi valori in base al proprio caso d'uso.

dimensione del broker	num.replica.fetchers
kafka.m5.4xlarge	4
kafka.m5.8xlarge	8
kafka.m5.12xlarge	14
kafka.m5.16xlarge	16
kafka.m5.24xlarge	16

La configurazione aggiornata potrebbe non avere effetto per un massimo di 24 ore e potrebbe richiedere più tempo quando un volume sorgente non è completamente utilizzato. Tuttavia, le prestazioni dei volumi di transizione sono almeno uguali a quelle dei volumi di archiviazione di origine durante il periodo di migrazione. Un volume da 1 TiB completamente utilizzato richiede in genere circa sei ore per migrare a una configurazione aggiornata.

Eseguire il provisioning della velocità di storage utilizzando AWS Management Console

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Scegli Create cluster (Crea cluster).
3. Scegli Creazione personalizzata.
4. Specificare un nome per il cluster.
5. Nella sezione Archiviazione, scegli Abilita.
6. Scegli un valore per la velocità di trasmissione effettiva dell'archiviazione per broker.
7. Scegli un VPC, zone e sottoreti, nonché un gruppo di sicurezza.
8. Seleziona Successivo.
9. Nella parte inferiore del passaggio Sicurezza, scegli Avanti.

10. Nella parte inferiore del passaggio Monitoraggio e tag, scegli Avanti.
11. Verifica le impostazioni del cluster, quindi scegli Crea cluster.

Eseguire il provisioning del throughput di storage utilizzando AWS CLI

Questa sezione mostra un esempio di come è possibile utilizzare il AWS CLI per creare un cluster con il throughput assegnato abilitato.

1. Copia il codice JSON seguente e incollalo in un file. Sostituisci i segnaposto degli ID di sottorete e gruppo di sicurezza con i valori del tuo account. Assegna al file il nome `cluster-creation.json` e salvalo.

```
{
    "Provisioned": {
        "BrokerNodeGroupInfo":{
            "InstanceType":"kafka.m5.4xlarge",
            "ClientSubnets":[
                "Subnet-1-ID",
                "Subnet-2-ID"
            ],
            "SecurityGroups":[
                "Security-Group-ID"
            ],
            "StorageInfo": {
                "EbsStorageInfo": {
                    "VolumeSize": 10,
                    "ProvisionedThroughput": {
                        "Enabled": true,
                        "VolumeThroughput": 250
                    }
                }
            }
        },
        "EncryptionInfo": {
            "EncryptionInTransit": {
                "InCluster": false,
                "ClientBroker": "PLAINTEXT"
            }
        },
        "KafkaVersion":"2.8.1",
        "NumberOfBrokerNodes": 2
    },
```

```
    "ClusterName": "provisioned-throughput-example"
}
```

2. Esegui il AWS CLI comando seguente dalla directory in cui hai salvato il file JSON nel passaggio precedente.

```
aws kafka create-cluster-v2 --cli-input-json file://cluster-creation.json
```

Assegnazione della velocità di trasmissione effettiva dell'archiviazione tramite l'API

Per configurare il throughput di storage assegnato durante la creazione di un cluster, usa V2. CreateCluster

Aggiornamento delle dimensioni del broker

È possibile scalare il cluster MSK su richiesta modificando le dimensioni dei broker senza riassegnare le partizioni di Apache Kafka. La modifica delle dimensioni dei broker offre la flessibilità necessaria per adattare la capacità di calcolo del cluster MSK in base alle variazioni dei carichi di lavoro, senza interrompere l'I/O del cluster. Amazon MSK utilizza le stesse dimensioni di broker per tutti i broker di un determinato cluster.

Questa sezione descrive come aggiornare le dimensioni del broker per il cluster MSK. È possibile aggiornare le dimensioni del broker del cluster da M5 o T3 a M7g o da M7g a M5. Tieni presente che la migrazione a un broker di dimensioni inferiori può ridurre le prestazioni e ridurre il throughput massimo raggiungibile per broker. La migrazione a un broker di dimensioni maggiori può aumentare le prestazioni ma può costare di più.

L'aggiornamento delle dimensioni di un broker avviene in modo continuativo mentre il cluster è attivo e funzionante. Ciò significa che Amazon MSK disattiva un broker alla volta per eseguire l'aggiornamento delle dimensioni del broker. Per informazioni su come rendere altamente disponibile un cluster durante un aggiornamento delle dimensioni di un broker, consulta. the section called “Creazione di cluster a disponibilità elevata” Per ridurre ulteriormente il potenziale impatto sulla produttività, è possibile eseguire l'aggiornamento delle dimensioni del broker durante un periodo di traffico ridotto.

Durante un aggiornamento delle dimensioni di un broker, puoi continuare a produrre e consumare dati. Tuttavia, è necessario attendere il completamento dell'aggiornamento prima di poter riavviare i broker o richiamare una delle operazioni di aggiornamento elencate nelle operazioni di Amazon MSK.

Se desideri aggiornare il cluster a un broker di dimensioni inferiori, ti consigliamo di provare prima l'aggiornamento su un cluster di test per vedere come influisce sullo scenario.

> **Important**
>
> Non puoi aggiornare un cluster a un broker di dimensioni inferiori se il numero di partizioni per broker supera il numero massimo specificato in. the section called “ Dimensionamento corretto del cluster: numero di partizioni per broker”

Aggiornamento delle dimensioni del broker utilizzando il AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegliete il cluster MSK per il quale desiderate aggiornare le dimensioni del broker.
3. Nella pagina dei dettagli del cluster, trova la sezione di riepilogo dei broker e scegli Modifica le dimensioni del broker.
4. Scegli la dimensione del broker che desideri dall'elenco.
5. Salva le modifiche.

Aggiornamento delle dimensioni del broker utilizzando il AWS CLI

1. Esegui il comando seguente, sostituendolo *ClusterArn*con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

   Sostituisci *Current-Cluster-Version* con la versione corrente del cluster e *TargetType*con la nuova dimensione che desideri che abbiano i broker. Per ulteriori informazioni sulle dimensioni dei broker, consulta. the section called “Dimensioni dei broker”

   ```
   aws kafka update-broker-type --cluster-arn ClusterArn --current-version Current-Cluster-Version --target-instance-type TargetType
   ```

   Di seguito è riportato un esempio di come utilizzare questo comando:

```
aws kafka update-broker-type --cluster-arn "arn:aws:kafka:us-
east-1:0123456789012:cluster/exampleName/abcd1234-0123-abcd-5678-1234abcd-1" --
current-version "K1X5R6FKA87" --target-instance-type kafka.m5.large
```

L'output di questo comando è simile all'esempio JSON seguente.

```
{
    "ClusterArn": "arn:aws:kafka:us-east-1:0123456789012:cluster/exampleName/
abcd1234-0123-abcd-5678-1234abcd-1",
    "ClusterOperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef"
}
```

2. Per ottenere il risultato dell'`update-broker-type`operazione, esegui il comando seguente, sostituendo *`ClusterOperationArn`* con l'ARN ottenuto nell'output del `update-broker-type` comando.

```
aws kafka describe-cluster-operation --cluster-operation-arn ClusterOperationArn
```

L'output di questo comando `describe-cluster-operation` è simile all'esempio JSON seguente.

```
{
  "ClusterOperationInfo": {
    "ClientRequestId": "982168a3-939f-11e9-8a62-538df00285db",
    "ClusterArn": "arn:aws:kafka:us-east-1:0123456789012:cluster/exampleName/
abcd1234-0123-abcd-5678-1234abcd-1",
    "CreationTime": "2021-01-09T02:24:22.198000+00:00",
    "OperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-operation/
exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef",
    "OperationState": "UPDATE_COMPLETE",
    "OperationType": "UPDATE_BROKER_TYPE",
    "SourceClusterInfo": {
      "InstanceType": "t3.small"
    },
    "TargetClusterInfo": {
      "InstanceType": "m5.large"
    }
```

```
  }
}
```

Se il valore di `OperationState` è `UPDATE_IN_PROGRESS`, attendi qualche minuto, quindi esegui nuovamente il comando `describe-cluster-operation`.

Aggiornamento delle dimensioni del broker tramite l'API

Per aggiornare le dimensioni del broker utilizzando l'API, vedi UpdateBrokerTipo.

Puoi utilizzarlo `UpdateBrokerType` per aggiornare le dimensioni del broker del cluster da M5 o T3 a M7g o da M7g a M5.

Aggiornamento della configurazione di un cluster Amazon MSK

Per aggiornare la configurazione di un cluster, assicurati che lo stato del cluster sia `ACTIVE`. Inoltre, devi assicurarti che il numero di partizioni per broker sul cluster MSK sia inferiore ai limiti descritti nella sezione the section called “ Dimensionamento corretto del cluster: numero di partizioni per broker”. Non è possibile aggiornare la configurazione di un cluster che supera questi limiti.

Per informazioni sulla configurazione MSK, incluso come creare una configurazione personalizzata, quali proprietà è possibile aggiornare e cosa accade quando si aggiorna la configurazione di un cluster esistente, consulta Configurazione.

Aggiornamento della configurazione di un cluster utilizzando AWS CLI

1. Copiare il JSON seguente e salvarlo in un file. Assegnare un nome al file `configuration-info.json`. Sostituisci *`ConfigurationArn`*con l'Amazon Resource Name (ARN) della configurazione che desideri utilizzare per aggiornare il cluster. La stringa ARN deve essere racchiusa tra virgolette nel seguente JSON.

   Sostituisci *`Configuration-Revision`* con la revisione della configurazione che desideri utilizzare. Le revisioni di configurazione sono interi (numeri interi) che iniziano da 1. Questo intero non deve essere racchiuso tra virgolette nel seguente JSON.

   ```
   {
        "Arn": ConfigurationArn,
        "Revision": Configuration-Revision
   }
   ```

2. Esegui il comando seguente, sostituendolo *ClusterArn*con l'ARN ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

   Sostituisci *Path-to-Config-Info-File* con il percorso del file delle informazioni di configurazione. Se il file creato nella fase precedente è stato denominato `configuration-info.json` e salvato nella directory corrente, *Path-to-Config-Info-File* è `configuration-info.json`.

   Sostituisci *Current-Cluster-Version* con la versione corrente del cluster.

   > **Important**
   >
   > Le versioni del cluster non sono interi semplici. Per trovare la versione corrente del cluster, usa l'DescribeClusteroperazione o il comando AWS CLI describe-cluster. Una versione di esempio è `KTVPDKIKX0DER`.

```
aws kafka update-cluster-configuration --cluster-arn ClusterArn --configuration-
info file://Path-to-Config-Info-File --current-version Current-Cluster-Version
```

   Di seguito è riportato un esempio di come utilizzare questo comando:

```
aws kafka update-cluster-configuration --cluster-arn "arn:aws:kafka:us-
east-1:0123456789012:cluster/exampleName/abcd1234-0123-abcd-5678-1234abcd-1" --
configuration-info file://c:\users\tester\msk\configuration-info.json --current-
version "K1X5R6FKA87"
```

   L'output di questo comando `update-cluster-configuration` è simile all'esempio JSON seguente.

```
{
    "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/exampleClusterName/
abcdefab-1234-abcd-5678-cdef0123ab01-2",
    "ClusterOperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef"
}
```

3. Per ottenere il risultato dell'`update-cluster-configuration`operazione, esegui il comando seguente, sostituendo *`ClusterOperationArn`* con l'ARN ottenuto nell'output del `update-cluster-configuration` comando.

```
aws kafka describe-cluster-operation --cluster-operation-arn ClusterOperationArn
```

L'output di questo comando `describe-cluster-operation` è simile all'esempio JSON seguente.

```
{
    "ClusterOperationInfo": {
        "ClientRequestId": "982168a3-939f-11e9-8a62-538df00285db",
        "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/
exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2",
        "CreationTime": "2019-06-20T21:08:57.735Z",
        "OperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef",
        "OperationState": "UPDATE_COMPLETE",
        "OperationType": "UPDATE_CLUSTER_CONFIGURATION",
        "SourceClusterInfo": {},
        "TargetClusterInfo": {
            "ConfigurationInfo": {
                "Arn": "arn:aws:kafka:us-east-1:123456789012:configuration/
ExampleConfigurationName/abcdabcd-abcd-1234-abcd-abcd123e8e8e-1",
                "Revision": 1
            }
        }
    }
}
```

In questo output, `OperationType` è `UPDATE_CLUSTER_CONFIGURATION`. Se il valore di `OperationState` è `UPDATE_IN_PROGRESS`, attendi qualche minuto, quindi esegui nuovamente il comando `describe-cluster-operation`.

Aggiornamento della configurazione di un cluster tramite l'API

Per utilizzare l'API per aggiornare la configurazione di un cluster, consulta UpdateCluster Configurazione.

Espansione di un cluster Amazon MSK

Utilizza questa operazione di Amazon MSK quando desideri incrementare il numero di broker nel cluster MSK. Per espandere un cluster, assicurati che il suo stato sia `ACTIVE`.

> **Important**
>
> Se desideri espandere un cluster MSK, assicurati di utilizzare questa operazione di Amazon MSK. Non provare ad aggiungere broker a un cluster senza utilizzare questa operazione.

Per informazioni su come ribilanciare le partizioni dopo aver aggiunto broker a un cluster, consulta the section called “Riassegnazione delle partizioni”.

Espansione di un cluster utilizzando AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli il cluster MSK di cui desideri aumentare numero di broker.
3. Nella pagina dei dettagli del cluster, scegli il pulsante Modifica accanto all'intestazione Dettagli broker a livello di cluster.
4. Inserisci il numero di broker di cui deve disporre il cluster per zona di disponibilità, quindi scegli Salva modifiche.

Espansione di un cluster utilizzando AWS CLI

1. Esegui il comando seguente, sostituendolo *`ClusterArn`*con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

   Sostituisci *`Current-Cluster-Version`* con la versione corrente del cluster.

   > **Important**
   >
   > Le versioni del cluster non sono interi semplici. Per trovare la versione corrente del cluster, usa l'DescribeClusteroperazione o il comando AWS CLI describe-cluster. Una versione di esempio è `KTVPDKIKX0DER`.

Il parametro *Target-Number-of-Brokers* rappresenta il numero totale di nodi broker di cui deve disporre il cluster al termine di questa operazione. Il valore specificato per *Target-Number-of-Brokers* deve essere un numero intero maggiore del numero corrente di broker nel cluster. Deve anche essere un multiplo del numero di zone di disponibilità.

```
aws kafka update-broker-count --cluster-arn ClusterArn --current-version Current-Cluster-Version --target-number-of-broker-nodes Target-Number-of-Brokers
```

L'output di questa operazione `update-broker-count` è simile al seguente JSON.

```
{

    "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2",
    "ClusterOperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-abcd-4f7f-1234-9876543210ef"
}
```

2. Per ottenere il risultato dell'`update-broker-count`operazione, esegui il comando seguente, sostituendo *ClusterOperationArn* con l'ARN ottenuto nell'output del `update-broker-count` comando.

```
aws kafka describe-cluster-operation --cluster-operation-arn ClusterOperationArn
```

L'output di questo comando `describe-cluster-operation` è simile all'esempio JSON seguente.

```
{
    "ClusterOperationInfo": {
        "ClientRequestId": "c0b7af47-8591-45b5-9c0c-909a1a2c99ea",
        "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2",
        "CreationTime": "2019-09-25T23:48:04.794Z",
        "OperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-abcd-4f7f-1234-9876543210ef",
        "OperationState": "UPDATE_COMPLETE",
        "OperationType": "INCREASE_BROKER_COUNT",
```

```
        "SourceClusterInfo": {
            "NumberOfBrokerNodes": 9
        },
        "TargetClusterInfo": {
            "NumberOfBrokerNodes": 12
        }
    }
}
```

In questo output, `OperationType` è `INCREASE_BROKER_COUNT`. Se il valore di `OperationState` è `UPDATE_IN_PROGRESS`, attendi qualche minuto, quindi esegui nuovamente il comando `describe-cluster-operation`.

Espansione di un cluster tramite l'API

Per aumentare il numero di broker in un cluster che utilizzano l'API, consulta Count. UpdateBroker

Rimuovere un broker da un cluster Amazon MSK

Usa questa operazione Amazon MSK quando desideri rimuovere broker dai cluster con provisioning di Amazon Managed Streaming for Apache Kafka (MSK). Puoi ridurre la capacità di storage e di elaborazione del cluster rimuovendo set di broker, senza alcun impatto sulla disponibilità, rischio di durabilità dei dati o interruzione delle applicazioni di streaming di dati.

Puoi aggiungere altri broker al cluster per gestire l'aumento del traffico e rimuovere i broker quando il traffico diminuisce. Grazie alla funzionalità di aggiunta e rimozione dei broker, è possibile utilizzare al meglio la capacità del cluster e ottimizzare i costi dell'infrastruttura MSK. La rimozione dei broker offre il controllo a livello di broker sulla capacità del cluster esistente per soddisfare le esigenze di carico di lavoro ed evitare la migrazione verso un altro cluster.

Utilizza la AWS console, l'interfaccia a riga di comando (CLI), l'SDK o AWS CloudFormation per ridurre il numero di broker del cluster a cui è stato assegnato il provisioning. MSK seleziona i broker che non dispongono di alcuna partizione (ad eccezione degli argomenti Canary) e impedisce alle applicazioni di produrre dati per tali broker, rimuovendo al contempo in modo sicuro tali broker dal cluster.

È necessario rimuovere un broker per zona di disponibilità, se si desidera ridurre lo storage e l'elaborazione di un cluster. Ad esempio, è possibile rimuovere due broker da un cluster con due

zone di disponibilità o tre broker da un cluster con tre zone di disponibilità in un'unica operazione di rimozione dei broker.

Per informazioni su come ribilanciare le partizioni dopo aver rimosso i broker da un cluster, vedere. the section called “Riassegnazione delle partizioni”

È possibile rimuovere i broker da tutti i cluster con provisioning MSK basati su M5 e M7g, indipendentemente dalla dimensione dell'istanza.

La rimozione dei broker è supportata nelle versioni di Kafka 2.8.1 e successive, inclusi i cluster in modalità Kraft.

Argomenti

Preparati a rimuovere i broker rimuovendo tutte le partizioni

Prima di iniziare il processo di rimozione del broker, sposta innanzitutto tutte le partizioni, tranne quelle relative agli argomenti `__amazon_msk_canary` e ai broker che `__amazon_msk_canary_state` intendi rimuovere. Si tratta di argomenti interni creati da Amazon MSK per i parametri diagnostici e di salute dei cluster.

Puoi utilizzare le API di amministrazione di Kafka o Cruise Control per spostare le partizioni su altri broker che intendi conservare nel cluster. Vedi Riassegnare le partizioni.

Procedura di esempio per rimuovere le partizioni

Questa sezione è un esempio di come rimuovere le partizioni dal broker che intendi rimuovere. Supponiamo di avere un cluster con 6 broker, 2 broker in ogni AZ e che abbia quattro argomenti:

- `__amazon_msk_canary`
- `__consumer_offsets`
- `__amazon_msk_connect_offsets_my-mskc-connector_12345678-09e7-c657f7e4ff32-2`
- `msk-brk-rmv`

1. Crea una macchina client come descritto in Creare una macchina client.
2. Dopo aver configurato il computer client, esegui il comando seguente per elencare tutti gli argomenti disponibili nel cluster.

```
./bin/kafka-topics.sh --bootstrap-server “CLUSTER_BOOTSTRAP_STRING” --list
```

   In questo esempio, vediamo quattro nomi di argomenti,`__amazon_msk_canary`, `__consumer_offsets__amazon_msk_connect_offsets_my-mskc-connector_12345678-09e7-c657f7e4ff32-2`, `emsk-brk-rmv`.
3. Crea un file json chiamato `topics.json` sul computer client e aggiungi tutti i nomi degli argomenti utente come nel seguente esempio di codice. Non è necessario includere il nome dell'`__amazon_msk_canary`argomento in quanto si tratta di un argomento gestito dal servizio che verrà spostato automaticamente quando necessario.

```
{
"topics": [
{"topic": "msk-brk-rmv"},
{"topic": "__consumer_offsets"},
{"topic": "__amazon_msk_connect_offsets_my-mskc-connector_12345678-09e7-
c657f7e4ff32-2"}
],
"version":1
}
```

4. Esegui il comando seguente per generare una proposta per spostare le partizioni su soli 3 broker su 6 broker del cluster.

```
./bin/kafka-reassign-partitions.sh --bootstrap-server “CLUSTER_BOOTSTRAP_STRING” --
topics-to-move-json-file topics.json --broker-list 1,2,3 --generate
```

5. Crea un file chiamato `reassignment-file.json` e copia il comando `proposed partition reassignment configuration` che hai ottenuto dal precedente comando.
6. Esegui il seguente comando per spostare le partizioni specificate in. `reassignment-file.json`

```
./bin/kafka-reassign-partitions.sh --bootstrap-server “CLUSTER_BOOTSTRAP_STRING” --
reassignment-json-file reassignment-file.json --execute
```

   L'esito si presenta in maniera analoga all'immagine riportata di seguito.

```
Successfully started partition reassignments for morpheus-test-topic-1-0,test-
topic-1-0
```

7. Esegui il comando seguente per verificare che tutte le partizioni siano state spostate.

```
./bin/kafka-reassign-partitions.sh --bootstrap-server “CLUSTER_BOOTSTRAP_STRING” --
reassignment-json-file reassignment-file.json --verify
```

   L'output è simile al seguente. Monitora lo stato fino a quando tutte le partizioni negli argomenti richiesti non sono state riassegnate correttamente:

```
Status of partition reassignment:
Reassignment of partition msk-brk-rmv-0 is completed.
Reassignment of partition msk-brk-rmv-1 is completed.
Reassignment of partition __consumer_offsets-0 is completed.
Reassignment of partition __consumer_offsets-1 is completed.
```

8. Quando lo stato indica che la riassegnazione delle partizioni per ogni partizione è stata completata, monitora le `UserPartitionExists` metriche per 5 minuti per assicurarti che vengano visualizzate dai broker da cui hai spostato le `0` partizioni. Dopo averlo confermato, puoi procedere alla rimozione del broker dal cluster.

Rimuovi un broker con la console di AWS gestione

Per rimuovere i broker con la console di gestione AWS

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli il cluster MSK che contiene i broker che desideri rimuovere.
3. Nella pagina dei dettagli del cluster, scegli il pulsante Azioni e seleziona l'opzione Modifica numero di broker.
4. Inserisci il numero di broker che desideri che il cluster abbia per zona di disponibilità. La console riepiloga il numero di broker nelle zone di disponibilità che verranno rimossi. Assicurati che sia quello che vuoi.
5. Seleziona Salvataggio delle modifiche.

Per evitare la rimozione accidentale del broker, la console ti chiede di confermare che desideri eliminare i broker.

Rimuovi un broker con la AWS CLI

Esegui il comando seguente, sostituendolo `ClusterArn` con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta Listing Amazon MSK clusters. Sostituisci `Current-Cluster-Version` con la versione corrente del cluster.

> **Important**
>
> Le versioni del cluster non sono interi semplici. Per trovare la versione corrente del cluster, usa l'DescribeClusteroperazione o il comando AWS CLI describe-cluster. Una versione di esempio è `KTVPDKIKX0DER`.

Il parametro *`Target-Number-of-Brokers`* rappresenta il numero totale di nodi broker di cui deve disporre il cluster al termine di questa operazione. Il valore specificato per *`Target-number-of-Brokers`* deve essere un numero intero inferiore al numero corrente di broker nel cluster. Deve anche essere un multiplo del numero di zone di disponibilità.

```
aws kafka update-broker-count --cluster-arn ClusterArn --current-version Current-
Cluster-Version --target-number-of-broker-nodes Target-Number-of-Brokers
```

L'output di questa operazione `update-broker-count` è simile al seguente JSON.

```
{
"ClusterOperationInfo": {
"ClientRequestId": "c0b7af47-8591-45b5-9c0c-909a1a2c99ea",
        "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/exampleClusterName/
abcdefab-1234-abcd-5678-cdef0123ab01-2",
        "CreationTime": "2019-09-25T23:48:04.794Z",
        "OperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef",
        "OperationState": "UPDATE_COMPLETE",
        "OperationType": "DECREASE_BROKER_COUNT",
        "SourceClusterInfo": {
"NumberOfBrokerNodes": 12
```

```
        },
        "TargetClusterInfo": {
"NumberOfBrokerNodes": 9
        }
    }
}
```

In questo output, `OperationType` è `DECREASE_BROKER_COUNT`. Se il valore di `OperationState` è `UPDATE_IN_PROGRESS`, attendi qualche minuto, quindi esegui nuovamente il comando `describe-cluster-operation`.

AWS Rimuovi un broker con l'API

Per rimuovere i broker in un cluster utilizzando l'API, consulta UpdateBrokerCount nell'Amazon Managed Streaming for Apache Kafka API Reference.

Aggiornamento delle impostazioni di sicurezza di un cluster

Utilizza questa operazione Amazon MSK per aggiornare le impostazioni di autenticazione e crittografia client-broker del tuo cluster MSK. Puoi anche aggiornare la Private Security Authority utilizzata per firmare i certificati per l'autenticazione TLS reciproca. Non è possibile modificare l'impostazione di crittografia all'interno del cluster (da broker a broker).

Per poter aggiornare le impostazioni di sicurezza, il cluster deve essere nello stato `ACTIVE`.

Se attivi l'autenticazione tramite IAM, SASL o TLS, devi attivare anche la crittografia tra client e broker. La tabella di seguito riporta le possibili combinazioni.

Autenticazione	Opzioni di crittografia client-broker	Crittografia broker-broker
Unauthenticated	TLS, PLAINTEXT, TLS_PLAINTEXT	Può essere attiva o non attiva.
mTLS	TLS, TLS_PLAINTEXT	Deve essere attiva.
SASL/SCRAM	TLS	Deve essere attiva.
SASL/IAM	TLS	Deve essere attiva.

Quando la crittografia client-broker è impostata su `TLS_PLAINTEXT` e l'autenticazione client è impostata su `mTLS`, Amazon MSK crea due tipi di ascoltatore a cui i client possono connettersi: un ascoltatore a cui i client possono connettersi utilizzando l'autenticazione mTLS con crittografia TLS e un altro a cui i client possono connettersi senza autenticazione o crittografia (non crittografato).

Per ulteriori informazioni sulle impostazioni di sicurezza, consulta la sezione Sicurezza.

Aggiornamento delle impostazioni di sicurezza di un cluster utilizzando AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Seleziona il cluster MSK che desideri aggiornare.
3. Nella sezione Impostazioni di sicurezza, scegli Modifica.
4. Scegli le impostazioni di autenticazione e crittografia che desideri per il cluster, quindi seleziona Salva modifiche.

Aggiornamento delle impostazioni di sicurezza di un cluster utilizzando AWS CLI

1. Crea un file JSON contenente le impostazioni di crittografia che desideri assegnare al cluster. Di seguito è riportato un esempio.

   > **Note**
   >
   > È possibile aggiornare solo l'impostazione di crittografia client-broker. Non è possibile aggiornare l'impostazione di crittografia all'interno del cluster (da broker a broker).

   ```
   {"EncryptionInTransit":{"ClientBroker": "TLS"}}
   ```

2. Crea un file JSON contenente le impostazioni di autenticazione che desideri che il cluster utilizzi. Di seguito è riportato un esempio.

   ```
   {"Sasl":{"Scram":{"Enabled":true}}}
   ```

3. Esegui il AWS CLI comando seguente:

```
aws kafka update-security --cluster-arn ClusterArn --current-version Current-
Cluster-Version --client-authentication file://Path-to-Authentication-Settings-
JSON-File --encryption-info file://Path-to-Encryption-Settings-JSON-File
```

L'output di questa operazione `update-security` è simile al seguente JSON.

```
{

    "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/exampleClusterName/
abcdefab-1234-abcd-5678-cdef0123ab01-2",
    "ClusterOperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef"
}
```

4. Per visualizzare lo stato dell'`update-security`operazione, esegui il comando seguente, sostituendo *`ClusterOperationArn`* con l'ARN ottenuto nell'output del `update-security` comando.

```
aws kafka describe-cluster-operation --cluster-operation-arn ClusterOperationArn
```

L'output di questo comando `describe-cluster-operation` è simile all'esempio JSON seguente.

```
{
    "ClusterOperationInfo": {
        "ClientRequestId": "c0b7af47-8591-45b5-9c0c-909a1a2c99ea",
        "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/
exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2",
        "CreationTime": "2021-09-17T02:35:47.753000+00:00",
        "OperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef",
        "OperationState": "PENDING",
        "OperationType": "UPDATE_SECURITY",
        "SourceClusterInfo": {},
        "TargetClusterInfo": {}
    }
}
```

Se il valore di `OperationState` è `PENDING` oppure `UPDATE_IN_PROGRESS`, attendi qualche minuto, quindi esegui nuovamente il comando `describe-cluster-operation`.

Aggiornamento delle impostazioni di sicurezza di un cluster tramite l'API

Per aggiornare le impostazioni di sicurezza per un cluster utilizzando l'API, vedere. UpdateSecurity

> **Note**
>
> Le operazioni AWS CLI e le API per l'aggiornamento delle impostazioni di sicurezza di un cluster sono idempotenti. Ciò significa che se si richiama l'operazione di aggiornamento della sicurezza e si specifica un'impostazione di autenticazione o crittografia uguale a quella attualmente utilizzata dal cluster, tale impostazione non verrà modificata.

Riavvio di un broker per un cluster Amazon MSK

Utilizza questa operazione di Amazon MSK quando desideri riavviare un broker per un cluster MSK. Per riavviare un broker per un cluster, assicurati che il cluster si trovi nello stato `ACTIVE`.

Il servizio Amazon MSK può riavviare i broker del cluster MSK durante la manutenzione del sistema, ad esempio durante l'applicazione di patch o gli aggiornamenti di versione. Il riavvio manuale di un broker consente di testare la resilienza dei client Kafka per determinare come rispondono alla manutenzione del sistema.

Riavvio di un broker utilizzando il AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli il cluster MSK di cui desideri riavviare il broker.
3. Scorri verso il basso fino alla sezione Dettagli del broker e scegli il broker che desideri riavviare.
4. Scegli il pulsante Riavvia broker.

Riavvio di un broker utilizzando il AWS CLI

1. Esegui il comando seguente, sostituendolo *ClusterArn* con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster e *BrokerId* poi con l'ID del broker che desideri riavviare.

   > **Note**
   >
   > L'operazione `reboot-broker` supporta il riavvio di un singolo broker alla volta.

   Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

   Se non disponi degli ID del broker per il tuo cluster, puoi trovarli elencando i nodi dei broker. Per ulteriori informazioni, consulta la sezione list-nodes.

   ```
   aws kafka reboot-broker --cluster-arn ClusterArn --broker-ids BrokerId
   ```

   L'output di questa operazione `reboot-broker` è simile al seguente JSON.

   ```
   {

       "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/exampleClusterName/
   abcdefab-1234-abcd-5678-cdef0123ab01-2",
       "ClusterOperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
   operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
   abcd-4f7f-1234-9876543210ef"
   }
   ```

2. Per ottenere il risultato dell'`reboot-broker`operazione, esegui il comando seguente, sostituendolo *ClusterOperationArn* con l'ARN ottenuto nell'output del `reboot-broker` comando.

   ```
   aws kafka describe-cluster-operation --cluster-operation-arn ClusterOperationArn
   ```

   L'output di questo comando `describe-cluster-operation` è simile all'esempio JSON seguente.

```
{
    "ClusterOperationInfo": {
        "ClientRequestId": "c0b7af47-8591-45b5-9c0c-909a1a2c99ea",
        "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/
exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2",
        "CreationTime": "2019-09-25T23:48:04.794Z",
        "OperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef",
        "OperationState": "REBOOT_IN_PROGRESS",
        "OperationType": "REBOOT_NODE",
        "SourceClusterInfo": {},
        "TargetClusterInfo": {}
    }
}
```

Quando l'operazione di riavvio è completa, il valore di `OperationState` è `REBOOT_COMPLETE`.

Riavvio di un broker tramite l'API

Per riavviare un broker in un cluster utilizzando l'API, vedere. RebootBroker

Impatto del riavvio del broker durante l'applicazione di patch e altre operazioni di manutenzione

Periodicamente, Amazon MSK aggiorna il software dei tuoi broker. Questi aggiornamenti non hanno alcun impatto sulle scritture e le letture delle applicazioni se segui le best practice.

Amazon MSK utilizza aggiornamenti periodici per il software per mantenere un'elevata disponibilità dei cluster. Durante questo processo, i broker vengono riavviati uno alla volta e Kafka trasferisce automaticamente la leadership a un altro broker online. I client Kafka dispongono di meccanismi integrati per rilevare automaticamente il cambio di leadership per le partizioni e continuare a scrivere e leggere dati in un cluster MSK.

In seguito alla disconnessione di un broker, è normale riscontrare errori transitori di disconnessione sui clienti. Inoltre, per una breve finestra (fino a 2 minuti, in genere meno), osserverete alcuni picchi nella latenza di lettura e scrittura di p99 (in genere alti millisecondi, fino a ~2 secondi). Questi picchi

sono previsti e sono causati dalla riconnessione del client a un nuovo broker leader; non influiscono sulla produzione o sul consumo e si risolveranno dopo la riconnessione.

Noterai anche un aumento della metrica`UnderReplicatedPartitions`, previsto in quanto le partizioni del broker che era stato chiuso non replicano più i dati. Ciò non ha alcun impatto sulle scritture e le letture delle applicazioni, poiché le repliche di queste partizioni ospitate su altri broker ora soddisfano le richieste.

Dopo l'aggiornamento del software, quando il broker torna online, deve «recuperare il ritardo» sui messaggi prodotti mentre era offline. Durante il catch up, si può anche osservare un aumento dell'utilizzo del volume, del throughput e della CPU. Questi non dovrebbero avere alcun impatto sulle scritture e le letture nel cluster se i broker dispongono di risorse sufficienti di CPU, memoria, rete e volume.

Assegnazione di tag a un cluster Amazon MSK

Puoi assegnare i tuoi metadati sotto forma di tag a una risorsa Amazon MSK, ad esempio un cluster MSK. Un tag è una coppia chiave-valore che definisci per la flusso. L'utilizzo dei tag è un modo semplice ma efficace per gestire AWS le risorse e organizzare i dati, inclusi i dati di fatturazione.

Argomenti

Nozioni di base sui tag

Puoi utilizzare l'API di Amazon MSK per completare le attività seguenti:

- Aggiunta di tag a una risorsa Amazon MSK.
- Elenco dei tag per una risorsa Amazon MSK.
- Rimozione dei tag da una risorsa Amazon MSK.

Puoi utilizzare i tag per categorizzare le risorse Amazon MSK. Ad esempio, puoi categorizzare i cluster Amazon MSK in base a scopo, proprietario o ambiente. Poiché definisci una chiave e un

valore per ogni tag, puoi creare un set di categorie personalizzate per soddisfare esigenze specifiche. Ad esempio, puoi definire un set di tag che consente di monitorare i cluster in base al proprietario e all'applicazione associata.

Di seguito sono illustrati alcuni esempi di tag:

- `Project: Project name`
- `Owner: Name`
- `Purpose: Load testing`
- `Environment: Production`

Monitoraggio dei costi mediante l'assegnazione di tag

Puoi utilizzare i tag per classificare e tenere traccia AWS dei costi. Quando applichi tag alle tue AWS risorse, inclusi i cluster Amazon MSK, il report sull'allocazione dei AWS costi include l'utilizzo e i costi aggregati per tag. Puoi organizzare i costi tra più servizi applicando tag che rappresentano categorie di business (come centri di costo, nomi di applicazioni o proprietari). Per ulteriori informazioni, consulta Utilizzo dei tag per l'allocazione dei costi ai fini dei report di fatturazione personalizzati nella AWS Billing User Guide (Guida per l'utente di Amazon API Gateway).

Limitazioni applicate ai tag

Ai tag in Amazon MSK si applicano le limitazioni seguenti.

Limitazioni di base

- Il numero massimo di tag per risorsa è 50.
- Per le chiavi e i valori dei tag viene fatta la distinzione tra maiuscole e minuscole.
- Non è possibile cambiare o modificare i tag di una risorsa eliminata.

Limitazioni applicate alle chiavi di tag

- Ogni chiave di tag deve essere univoca. Se aggiungi un tag con una chiave già in uso, il nuovo tag sovrascrive la coppia chiave-valore esistente.
- Una chiave di tag non può iniziare con `aws:` perché questo prefisso è riservato per l'utilizzo da parte di AWS. AWS crea tag con questo prefisso per tuo conto, ma non puoi modificarli o eliminarli.
- Le chiavi di tag devono avere una lunghezza compresa tra 1 e 128 caratteri Unicode.

- Le chiavi di tag devono contenere i seguenti caratteri: lettere Unicode, cifre, spazio e i seguenti caratteri speciali: `_ . / = + - @`.

Limitazioni applicate ai valori dei tag

- I valori dei tag devono avere una lunghezza compresa tra 0 e 255 caratteri Unicode.
- I valori dei tag possono essere vuoti. In caso contrario, devono contenere i seguenti caratteri: lettere Unicode, cifre, spazio e i seguenti caratteri speciali: `_ . / = + - @`.

Assegnazione di tag alle risorse tramite l'API di Amazon MSK

Puoi utilizzare le operazioni seguenti per assegnare o rimuovere i tag da una risorsa Amazon MSK o per elencare il set di tag corrente per una risorsa:

- ListTagsForResource
- TagResource
- UntagResource

Configurazione di Amazon MQ

Amazon Managed Streaming for Apache Kafka fornisce una configurazione predefinita per broker, argomenti e nodi Apache. ZooKeeper Puoi inoltre creare configurazioni personalizzate e utilizzarle per creare nuovi cluster MSK o per aggiornare cluster esistenti. Una configurazione MSK è costituita da un insieme di proprietà e dai relativi valori corrispondenti.

Argomenti

Configurazioni MSK personalizzate

Puoi utilizzare Amazon MSK per creare una configurazione MSK personalizzata in cui vengono impostate le proprietà seguenti. Le proprietà che non vengono impostate in modo esplicito ricevono i valori che hanno in the section called “Configurazione di default”. Per ulteriori informazioni sulle proprietà di configurazione, consulta la pagina relativa alla configurazione di Apache Kafka.

Proprietà di configurazione di Apache Kafka

Nome	Descrizione
allow.everyone.if.no.acl.found	Se desideri impostare questa proprietà su `false`, assicurati innanzitutto di definire le ACL di Apache Kafka per il cluster. Se imposti questa proprietà su `false` e non definisci prima le ACL di Apache Kafka, perderai l'accesso al cluster. In tal caso, puoi aggiornare nuovamente la configurazione e impostare questa proprietà su `true` per ottenere nuovamente l'accesso al cluster.
auto.create.topics.enable	Abilita la creazione automatica di argomenti sul server.

Nome	Descrizione
compression.type	Il tipo di compressione finale per un determinato argomento. Puoi impostare questa proprietà sui codec di compressione standard (`gzip`, `snappy`, `lz4` e `zstd`). Inoltre, accetta `uncompressed` . Questo valore equivale a nessuna compressione. Se imposti il valore su `producer`, significa mantenere il codec di compressione originale impostato dal produttore.
connections.max.idle.ms	Timeout delle connessioni inattive in millisecondi. I thread del processore dei socket del server chiudono le connessioni inattive per un periodo superiore al valore impostato per questa proprietà.
default.replication.factor	Il fattore di replica predefinito per argomenti creati automaticamente.
delete.topic.enable	Abilita l'operazione di eliminazione argomento . Se questa configurazione è disattivata, non è possibile eliminare un argomento tramite lo strumento di amministrazione.
group.initial.rebalance.delay.ms	Il tempo durante il quale il coordinatore del gruppo attende che altri consumatori si uniscano a un nuovo gruppo prima di eseguire il primo ribilanciamento. Un ritardo più lungo significa potenzialmente meno ribilanciamenti, ma aumenta il tempo prima dell'inizio dell'elaborazione.

Nome	Descrizione
group.max.session.timeout.ms	Timeout sessione massimo per i consumatori registrati. Timeout più lunghi offrono ai consumatori più tempo per elaborare i messaggi tra heartbeat, ma implicano un aumento del tempo richiesto per rilevare gli errori.
group.min.session.timeout.ms	Timeout sessione minimo per consumatori registrati. Timeout più brevi comportano un rilevamento più rapido degli errori, ma implicano un heartbeat dei consumatori più frequente. Ciò può sovraccaricare le risorse del broker.
leader.imbalance.per.broker.percentage	Il rapporto di squilibrio leader consentito per broker. Il controller attiva un bilanciamento dei leader se supera questo valore per broker. Questo valore è specificato in percentuale.
log.cleaner.delete.retention.ms	Quantità di tempo per cui Apache Kafka deve conservare i record eliminati. Il valore minimo è 0.

Nome	Descrizione
log.cleaner.min.cleanable.ratio	Questa proprietà di configurazione può avere valori compresi tra 0 e 1. Questo valore determina la frequenza con cui il compattatore di log tenta di pulire il log (se la compattazione dei log è abilitata). Per impostazione predefini ta, Apache Kafka evita di pulire un log se più del 50% del log è stato compattato. Questo rapporto limita lo spazio massimo che il log spreca con i duplicati (al 50%, ciò significa che al massimo il 50% del log potrebbe essere duplicato). Un rapporto più elevato significa un numero inferiore, pulizie più efficienti, ma anche più spreco di spazio nel log.
log.cleanup.policy	La policy di pulizia predefinita per i segmenti oltre la finestra di conservazione. Un elenco separato da virgole di policy valide. Policy valide sono `delete` e `compact`. Per i cluster abilitati all'archiviazione a più livelli, è valida solo la policy `delete`.
log.flush.interval.messages	Numero di messaggi che si accumulano in una partizione di log prima che i messaggi vengano scaricati su disco.
log.flush.interval.ms	Tempo massimo, in millisecondi, di mantenime nto in memoria di un messaggio in un argomento prima che venga scaricato su disco. Se questo valore non viene impostato, viene utilizzato il valore in log.flush.scheduler.interva l.ms. Il valore minimo è 0.

Nome	Descrizione
log.message.timestamp.difference.max.ms	La differenza massima consentita tra il timestamp del momento in cui un broker riceve un messaggio e il timestamp specificato nel messaggio. Se log.message.timestamp.type= , un messaggio viene rifiutato se la differenza di timestamp supera questa soglia. CreateTim e Questa configurazione viene LogAppend ignorata se log.message.timestamp.type= Time.
log.message.timestamp.type	Specifica se il timestamp nel messaggio è l'ora di creazione del messaggio o l'ora di aggiunta del log. I valori consentiti sono `CreateTime` e `LogAppendTime` .
log.retention.bytes	Dimensione massima del log prima dell'elim inazione.
log.retention.hours	Numero di ore per cui mantenere un file di log prima di eliminarlo, terziario alla proprietà log.retention.ms.
log.retention.minutes	Numero di minuti per cui mantenere un file di log prima di eliminarlo, secondario alla proprietà log.retention.ms. Se questo valore non viene impostato, viene utilizzato il valore in log.retention.hours.
log.retention.ms	Numero di millisecondi per cui mantenere un file di log prima di eliminarlo. Se non è impostato, viene utilizzato il valore in log.reten tion.minutes.

Nome	Descrizione
log.roll.ms	Tempo massimo prima che un nuovo segmento di log venga distribuito (in millisecondi). Se questo valore non viene impostato, viene utilizzato il valore in log.roll.hours. Il valore minimo possibile per questa proprietà è 1.
log.segment.bytes	Dimensione massima di un singolo file di log.
max.incremental.fetch.session.cache.slots	Numero massimo di sessioni di recupero incrementali che vengono mantenute.
message.max.bytes	Dimensione massima del batch di record consentita da Kafka. Se aumenti questo valore e sono presenti consumatori più vecchi di 0.10.2, anche le dimensioni di recupero dei consumatori devono essere incrementate in modo che possano recuperare batch di record di queste dimensioni. Nella versione più recente del formato del messaggio, i messaggi vengono sempre raggruppati in batch per maggiore efficienz a. Nelle versioni precedenti del formato del messaggio, i record non compressi non sono raggruppati in batch; in tal caso, questo limite si applica solo a un singolo record. Questo valore può essere impostato a livello di argomento con la configurazione max.messa ge.bytes.

Nome	Descrizione
min.insync.replicas	Quando un produttore imposta le ACK su `"all"` (o `"-1"`), il valore in min.insync.replica s specifica il numero minimo di repliche che devono riconoscere una scrittura affinché questa sia considerata correttamente completat a. Se questo minimo non può essere raggiunto , il produttore solleva un'eccezione (o). NotEnoughReplicas NotEnoughReplicasA fterAppend È possibile utilizzare i valori in min.insyn c.replicas e ACK per applicare maggiori garanzie di durabilità. Ad esempio, è possibile creare un argomento con un fattore di replica di 3, impostare min.insync.replicas su 2 e produrre con ACK di `"all"`. Ciò garantisce che il produttore generi un'eccezione se la maggior parte delle repliche non riceve una scrittura.
num.io.thread	Il numero di thread utilizzati dal server per elaborare le richieste, che possono includere I/ O del disco.
num.network.threads	Il numero di thread utilizzati dal server per ricevere richieste dalla rete e inviarle le risposte.
num.partitions	Numero predefinito di partizioni di log per argomento.
num.recovery.threads.per.data.dir	Il numero di thread per directory di dati da utilizzare per il ripristino dei log all'avvio e per lo scaricamento all'arresto.

Nome	Descrizione
num.replica.fetchers	Il numero di thread fetcher utilizzati per rispondere ai messaggi da un broker di origine. Incrementando questo valore, è possibile aumentare il grado di parallelismo I/O nel broker follower.
offsets.retention.minutes	Dopo che un gruppo di consumatori perde tutti i suoi consumatori (ovvero, diventa vuoto) i suoi offset vengono mantenuti per questo periodo di conservazione prima di essere scartati. Per i consumatori autonomi (ossia che utilizzano l'assegnazione manuale), gli offset scadono dopo l'ora dell'ultimo commit più questo periodo di conservazione.
offsets.topic.replication.factor	Il fattore di replica per l'argomento di offset. La selezione di un valore più alto garantisce la disponibilità. La creazione di argomenti interni non riesce fino a quando la dimensione del cluster non soddisfa questo requisito del fattore di replica.
replica.fetch.max.bytes	Numero di byte di messaggi da recuperare per ogni partizione. Questo valore non è un massimo assoluto. Se il primo batch di record nella prima partizione non vuota del recupero è più grande di questo valore, viene restituito il batch di record per garantire l'avanzamento. La proprietà message.max.bytes (configurazione broker) o max.message.bytes (configurazione argomento) specifica la dimensione massima del batch di record accettata dal broker.

Nome	Descrizione
replica.fetch.response.max.bytes	Il numero massimo di byte previsto per l'intera risposta di recupero. I record vengono recuperati in batch. Se il primo batch di record nella prima partizione non vuota del recupero è più grande di questo valore, il batch di record verrà comunque restituito per garantire l'avanzamento. Questo non è un massimo assoluto. Le proprietà message.max.bytes (configurazione broker) o max.message.bytes (configurazione argomento) specificano la dimensione massima del batch di record accettata dal broker.
replica.lag.time.max.ms	Se un follower non ha inviato richieste di fetch o non ha consumato fino all'offset di fine log del leader per almeno questo numero di milliseco ndi, il leader rimuove il follower dall'ISR. MinValue: 10000 MaxValue = 30000
replica.selector.class	Il nome completo della classe che implement a. ReplicaSelector Il broker utilizza questo valore per trovare la replica di lettura preferita . Se utilizzi Apache Kafka versione 2.4.1 o superiore e desideri consentire ai consumato ri di recuperare dati dalla replica più vicina, imposta questa proprietà su `org.apach e.kafka.common.replica.Rack AwareReplicaSelector` . Per ulteriori informazioni, consulta the section called “Apache Kafka versione 2.4.1 (usa invece 2.4.1.1)”.

Nome	Descrizione
replica.socket.receive.buffer.bytes	Il buffer di ricezione socket per le richieste di rete.
socket.receive.buffer.bytes	Buffer SO_RCVBUF dei socket del server dei socket. Il valore minimo che è possibile impostare per questa proprietà è -1. Se il valore è -1, Amazon MSK utilizza il sistema operativo predefinito.
socket.request.max.bytes	Il numero massimo di byte in una richiesta socket.
socket.send.buffer.bytes	Buffer SO_SNDBUF dei socket del server dei socket. Il valore minimo che è possibile impostare per questa proprietà è -1. Se il valore è -1, Amazon MSK utilizza il sistema operativo predefinito.
transaction.max.timeout.ms	Timeout massimo per transazioni. Se il tempo di transazione richiesto da un cliente supera questo valore, il broker restituisce un errore in. InitProducerIdRequest Ciò evita un timeout troppo elevato per un client, che può rallentar e i consumatori che leggono dagli argomenti inclusi nella transazione.
transaction.state.log.min.isr	Configurazione min.insync.replicas ignorata per l'argomento di transazione.
transaction.state.log.replication.factor	Il fattore di replica per l'argomento di transazio ne. La selezione di un valore più alto per questa proprietà aumenta la disponibilità. La creazione di argomenti interni non riesce fino a quando la dimensione del cluster non soddisfa questo requisito del fattore di replica.

Nome	Descrizione
transactional.id.expiration.ms	Il tempo in millisecondi durante il quale il coordinatore della transazione attende di ricevere eventuali aggiornamenti sullo stato delle transazioni per la transazione corrente prima che il coordinatore faccia scadere il proprio ID transazionale. Questa impostazione influenza anche la scadenza dell'ID produttore, poiché fa scadere gli ID produttore allo scadere di questo periodo di tempo dopo l'ultima scrittura con l'ID produttore specificato. Gli ID produttore potrebbero scadere prima se l'ultima scrittura dall'ID produttore viene eliminata a causa delle impostazioni di conservazione dell'argomento. Il valore minimo per questa proprietà è 1 millisecondo.
unclean.leader.election.enable	Indica se le repliche non incluse nel set ISR devono fungere da leader come ultima risorsa, anche se ciò potrebbe comportare la perdita di dati.
zookeeper.connection.timeout.ms	ZooKeeper cluster di modalità. Tempo massimo di attesa del client per stabilire una connessione. ZooKeeper Se questo valore non viene impostato, viene utilizzato il valore fornito in zookeeper.session.timeout.ms. MinValue = 6000 MaxValue (incluso) = 18000

Nome	Descrizione
zookeeper.session.timeout.ms	ZooKeeper più cluster. Il timeout della ZooKeeper sessione Apache in millisecondi. MinValue = 6000 MaxValue (incluso) = 18000

Per informazioni su come creare una configurazione MSK personalizzata, elencare tutte le configurazioni o descriverle, consulta the section called “Operazioni di configurazione”. Per creare un cluster MSK utilizzando una configurazione MSK personalizzata o per aggiornare un cluster con una nuova configurazione personalizzata, consulta la pagina Come funziona.

Quando si aggiorna il cluster MSK esistente con una configurazione MSK personalizzata, Amazon MSK esegue riavvii in sequenza quando necessario e utilizza le best practice per ridurre al minimo i tempi di inattività del cliente. Ad esempio, dopo aver riavviato ogni broker, Amazon MSK prova a lasciare che il broker recuperi i dati che potrebbe aver perso durante l'aggiornamento della configurazione prima di passare al broker successivo.

Configurazione dinamica

Oltre alle proprietà di configurazione fornite da Amazon MSK, puoi impostare dinamicamente le proprietà di configurazione a livello di cluster e broker che non richiedono un riavvio del broker. È possibile impostare dinamicamente alcune proprietà di configurazione. Si tratta delle proprietà che non sono contrassegnate come di sola lettura nella tabella in Broker Configs nella documentazione di Apache Kafka. Per informazioni sulla configurazione dinamica e sui comandi di esempio, consulta la pagina Updating Broker Configs nella documentazione di Apache Kafka.

> **Note**
>
> Puoi impostare la proprietà `advertised.listeners`, ma non la proprietà `listeners`.

Configurazione a livello di argomento

Puoi utilizzare i comandi Apache Kafka per impostare o modificare le proprietà di configurazione a livello dell'argomento per argomenti nuovi ed esistenti. Per ulteriori informazioni sulle proprietà di

configurazione a livello di argomento ed esempi di come impostarle, consulta la pagina Topic-Level Configs nella documentazione ufficiale di Apache Kafka.

Stati di configurazione

Una configurazione Amazon MSK può trovarsi in uno dei seguenti stati. Per eseguire un'operazione su una configurazione, la configurazione deve trovarsi nello stato `ACTIVE` o `DELETE_FAILED`:

- `ACTIVE`
- `DELETING`
- `DELETE_FAILED`

La configurazione predefinita di Amazon MSK

Quando crei un cluster MSK senza specificare una configurazione MSK personalizzata, Amazon MSK crea e utilizza una configurazione predefinita con i valori visualizzati nella tabella seguente. Per le proprietà non presenti in questa tabella, Amazon MSK utilizza i valori predefiniti associati alla versione di Apache Kafka. Per un elenco di questi valori predefiniti, consulta la pagina relativa alla configurazione di Apache Kafka.

Valori di configurazione predefiniti

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
allow.everyone.if.no.acl.found	Se nessun modello di risorse corrisponde a una risorsa specifica, la risorsa non dispone di ACL associati. In questo caso, se questa proprietà è impostata su `true`, tutti possono	`true`	`true`

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
	accedere alla risorsa, non solo i superutenti.		
auto.create.topics.enable	Abilita la creazione automatica di un argomento sul server.	`false`	`false`
auto.leader.rebalance.enable	Consente il bilanciamento leader automatico. Un thread in background controlla e attiva il bilanciamento del leader a intervalli regolari, se necessario.	`true`	`true`
default.replication.factor	Fattori di replica predefiniti per argomenti creati automaticamente.	3 per i cluster in 3 zone di disponibilità e 2 per i cluster in 2 zone di disponibilità.	3 per i cluster in 3 zone di disponibilità e 2 per i cluster in 2 zone di disponibilità.

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
local.retention.bytes	La dimensione massima dei segmenti di log locali per una partizione prima dell'eliminazione dei vecchi segmenti. Se questo valore non viene impostato, viene utilizzato il valore in log.retention.bytes. Il valore effettivo deve essere sempre minore o uguale al valore di log.retention.bytes. Il valore predefinito -2 indica che non è previsto un limite alla conservazione locale. Ciò corrisponde all'impostazione retention.ms/bytes di -1. Le proprietà local.retention.ms e local.retention.bytes sono simili a log.retention, in quanto vengono utilizzate per determinare per quanto tempo i segmenti di log devono rimanere	-2 per illimitato	-2 per illimitato

Nome	Descrizione	Valore predefinito per il cluster con l'archivi azione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
	nell'archiviazione locale. Le configura zioni log.reten tion.* esistenti sono configurazioni di conservazione per la partizione degli argomenti. Ciò include l'archiviazione locale e remota. Valori validi: numeri interi in [-2; +Inf]		

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
local.retention.ms	Numero di millisecondi di conservazione del segmento di log locale prima dell'eliminazione. Se questo valore non viene impostato, Amazon MSK utilizza il valore in log.retention.ms. Il valore effettivo deve essere sempre minore o uguale al valore di log.retention.bytes. Il valore predefinito -2 indica che non è previsto un limite alla conservazione locale. Ciò corrisponde all'impostazione retention.ms/bytes di -1. I valori local.retention.ms e local.retention.bytes sono simili a log.retention. MSK utilizza questa configurazione per determinare per quanto tempo i segmenti di log devono rimanere	-2 per illimitato	-2 per illimitato

Nome	Descrizione	Valore predefinito per il cluster con l'archivi azione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
	nell'archiviazione locale. Le configura zioni log.reten tion.* esistenti sono configurazioni di conservazione per la partizione degli argomenti. Ciò include l'archiviazione locale e remota. I valori validi sono numeri interi maggiori di 0.		

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
log.message.timestamp.difference.max.ms	La differenza massima consentita tra il timestamp quando un broker riceve un messaggio e il timestamp specificato nel messaggio. Se log.message.timestamp.type=CreateTime, un messaggio verrà rifiutato se la differenza di timestamp supera questa soglia. Questa configurazione viene ignorata se log.message.timestamp.type=LogAppendTime. La differenza di timestamp massima consentita non deve essere maggiore di log.retention.ms per evitare una distribuzione dei log inutilmente frequente.	9223372036854775807	86400000 per Kafka 2.8.2.tiered
log.segment.bytes	Dimensione massima di un singolo file di log.	1073741824	134217728

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
min.insync.replicas	Quando un produttore imposta il valore delle ACK (il riconoscimento che il produttore riceve dal broker Kafka) su `"all"` (o `"-1"`), il valore in min.insync.replicas specifica il numero minimo di repliche che devono riconoscere una scrittura affinché questa sia considerata correttamente completata. Se questo valore non soddisfa questo minimo, il produttore solleva un'eccezione (o). NotEnoughReplicas NotEnoughReplicasAfterAppend Se usati insieme, i valori min.insync.replicas e ACK consentono di imporre maggiori garanzie di durata. Ad esempio, è possibile creare un argomento con un fattore di replica di 3,	2 per i cluster in 3 zone di disponibilità e 1 per i cluster in 2 zone di disponibilità.	2 per i cluster in 3 zone di disponibilità e 1 per i cluster in 2 zone di disponibilità.

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
	impostare min.insync.replicas su 2 e produrre con ACK di `"all"`. Ciò garantisce che il produttore generi un'eccezione se la maggior parte delle repliche non riceve una scrittura.		
num.io.thread	Numero di thread utilizzati dal server per produrre le richieste, che possono includere l'I/O del disco.	8	max(8, vCPU), dove le vCPU dipendono dalla dimensione dell'istanza del broker
num.network.threads	Il numero di thread utilizzati dal server per ricevere richieste dalla rete e inviarle le risposte.	5	max(5, vCPU/2), dove le vCPU dipendono dalla dimensione dell'istanza del broker
num.partitions	Numero predefinito di partizioni di log per argomento.	1	1

Nome	Descrizione	Valore predefinito per il cluster con l'archivi azione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
num.replica.fetchers	Il numero di thread di recupero utilizzati per replicare i messaggi da un broker di origine. Se si aumenta questo valore, è possibile aumentare il grado di paralleli smo I/O nel broker follower.	2	max(2, vCPU/4), dove le vCPU dipendono dalla dimensione dell'istanza del broker
remote.log.msk.dis able.policy	Utilizzato con remote.storage.ena ble per disabilitare l'archiviazione a più livelli. Imposta questa policy su Elimina per indicare che i dati nell'archiviazione a più livelli vengono eliminati quando si imposta remote.st orage.enable su false.	N/D	DELETE
remote.log.reader. threads	La dimensione del pool di thread del lettore di log remoto, utilizzato nella pianificazione delle attività per recuperare dati dall'archiviazione remota.	N/D	max(10, vCPU*0,67), dove le vCPU dipendono dalla dimensione dell'ista nza del broker

Nome	Descrizione	Valore predefinito per il cluster con l'archivi azione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
remote.storage.ena ble	Abilita l'archiviazione a più livelli (remota) per un argomento, se impostato su true. Disabilita l'archivi azione a più livelli a livello di argomento se impostato su false e remote.lo g.msk.disable.policy è impostato su Delete. Quando si disabilita l'archiviazione a più livelli, si eliminano i dati dall'archiviazione remota. Una volta disabilitata l'archivi azione a più livelli su un argomento, non sarà possibile riabilita rla.	false	true

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
replica.lag.time.max.ms	Se un follower non ha inviato richieste di fetch o non ha consumato fino all'offset di fine log del leader per almeno questo numero di millisecondi, il leader rimuove il follower dall'ISR.	30000	30000

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
retention.ms	Campo obbligatorio. Il tempo minimo è 3 giorni. Non esiste un'impostazione predefinita perché l'impostazione è obbligatoria. Amazon MSK utilizza il valore retention.ms con local.retention.ms per determinare quando i dati vengono spostati dall'archiviazione locale a quella a più livelli. Il valore local.retention.ms specifica quando spostare i dati dall'archiviazione locale a quella a più livelli. Il valore retention.ms specifica quando rimuovere i dati dall'archiviazione a più livelli (ossia, rimuoverli dal cluster). Valori validi: numeri interi in [-1; +Inf]	Minimo 259.200.000 millisecondi (3 giorni). -1 per una conservazione infinita.	Minimo 259.200.000 millisecondi (3 giorni). -1 per una conservazione infinita.

Nome	Descrizione	Valore predefinito per il cluster con l'archiviazione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
socket.receive.buffer.bytes	Buffer SO_RCVBUF dei socket del server dei socket. Se il valore è -1, viene utilizzato il sistema operativo predefinito.	102400	102400
socket.request.max.bytes	Numero massimo di byte in una richiesta socket.	104857600	104857600
socket.send.buffer.bytes	Buffer SO_SNDBUF dei socket del server dei socket. Se il valore è -1, viene utilizzato il sistema operativo predefinito.	102400	102400
unclean.leader.election.enable	Indica se desideri che le repliche non incluse nel set ISR fungano da leader come ultima risorsa, anche se ciò potrebbe comportare la perdita di dati.	true	false
zookeeper.session.timeout.ms	Il timeout della ZooKeeper sessione Apache in millisecondi.	18000	18000

Nome	Descrizione	Valore predefinito per il cluster con l'archivi azione non a più livelli	Valore predefinito per il cluster abilitato all'archiviazione a più livelli
zookeeper.set.acl	Il client impostato per l'utilizzo di ACL sicure.	false	false

Per informazioni su come specificare valori di configurazione personalizzati, consulta la pagina the section called “Configurazioni personalizzate di ”.

Linee guida per la configurazione a livello di argomento dell'archiviazione a più livelli

Di seguito sono riportate le impostazioni e le limitazioni predefinite per la configurazione dell'archiviazione a più livelli a livello di argomento.

- Amazon MSK non supporta segmenti di log di dimensioni inferiori per argomenti con l'archiviazione a più livelli attivata. Se si desidera creare un segmento, è prevista una dimensione minima del segmento di log di 48 MiB o un tempo minimo di distribuzione del segmento di 10 minuti. Questi valori sono mappati alle proprietà segment.bytes e segment.ms.
- Il valore di local.retention.ms/bytes non può essere uguale o superiore a retention.ms/bytes. Questa è l'impostazione di conservazione dell'archiviazione a più livelli.
- Il valore predefinito per local.retention.ms/bytes è -2. Ciò significa che il valore retention.ms viene utilizzato per local.retention.ms/bytes. In questo caso, i dati rimangono sia nell'archiviazione locale sia nell'archiviazione a più livelli (una copia per ciascuna) e scadono insieme. Con questa opzione, una copia dei dati locali viene memorizzata nell'archiviazione remota. In questo caso, i dati letti dal traffico di utilizzo provengono dall'archiviazione locale.
- Il valore predefinito per retention.ms è 7 giorni. Non esiste un limite di dimensione predefinito per retention.bytes.
- Il valore minimo per retention.ms/bytes è -1. Ciò significa una conservazione infinita.
- Il valore minimo per local.retention.ms/bytes è -2. Ciò significa una conservazione infinita per l'archiviazione locale. Corrisponde all'impostazione di retention.ms/bytes su -1.

- La configurazione a livello di argomento retention.ms è obbligatoria per gli argomenti con l'archiviazione a più livelli attivata. Il valore minimo per retention.ms è 3 giorni.

Operazioni di configurazione di Amazon MSK

In questo argomento viene descritto come creare configurazioni MSK personalizzate e come eseguire operazioni su di esse. Per informazioni su come utilizzare configurazioni MSK per creare o aggiornare cluster, consulta Come funziona.

Questo argomento contiene le sezioni seguenti:

Per creare una configurazione MSK

1. Creare un file in cui specificare le proprietà di configurazione che si desidera impostare e i valori da assegnare alle stesse. Di seguito sono riportati i contenuti di un file di configurazione di esempio.

```
auto.create.topics.enable = true

log.roll.ms = 604800000
```

2. Esegui il AWS CLI comando seguente e sostituisci *config-file-path con il percorso* del file in cui hai salvato la configurazione nel passaggio precedente.

> **Note**
>
> Il nome scelto per la configurazione deve corrispondere alla seguente espressione regolare: "^[0-9A-Za-z][0-9A-Za-z-]{0,}$".

```
aws kafka create-configuration --name "ExampleConfigurationName" --description
 "Example configuration description." --kafka-versions "1.1.1" --server-properties
 fileb://config-file-path
```

Di seguito è riportato un esempio di una risposta corretta dopo l'esecuzione di questo comando.

```
{
    "Arn": "arn:aws:kafka:us-east-1:123456789012:configuration/SomeTest/
abcdabcd-1234-abcd-1234-abcd123e8e8e-1",
    "CreationTime": "2019-05-21T19:37:40.626Z",
    "LatestRevision": {
        "CreationTime": "2019-05-21T19:37:40.626Z",
        "Description": "Example configuration description.",
        "Revision": 1
    },
    "Name": "ExampleConfigurationName"
}
```

3. Il comando precedente restituisce un nome della risorsa Amazon (ARN) per la configurazione appena creata. Salvare questo ARN perché occorre per fare riferimento a questa configurazione in altri comandi. Se l'ARN della configurazione viene perso, è possibile trovarlo nuovamente elencando tutte le configurazioni presenti nell'account.

Aggiornamento di una configurazione MSK

1. Crea un file in cui specificare le proprietà di configurazione che desideri aggiornare e i valori da assegnare alle stesse. Di seguito sono riportati i contenuti di un file di configurazione di esempio.

```
auto.create.topics.enable = true

min.insync.replicas = 2
```

2. Esegui il comando AWS CLI , sostituendo *config-file-path* con il percorso del file in cui è stata salvata la configurazione nel passaggio precedente.

   Sostituisci *configuration-arn* con l'ARN ottenuto al momento della creazione della configurazione. Se l'ARN non è stato salvato al momento della creazione della configurazione, è possibile utilizzare il comando `list-configurations` per elencare tutte le configurazioni

presenti nell'account. La configurazione desiderata viene visualizzata nell'elenco di risposta. L'ARN della configurazione viene visualizzato anche in tale elenco.

```
aws kafka update-configuration --arn configuration-arn --description "Example
 configuration revision description." --server-properties fileb://config-file-path
```

3. Di seguito è riportato un esempio di una risposta corretta dopo l'esecuzione di questo comando.

```
{
    "Arn": "arn:aws:kafka:us-east-1:123456789012:configuration/SomeTest/
abcdabcd-1234-abcd-1234-abcd123e8e8e-1",
    "LatestRevision": {
        "CreationTime": "2020-08-27T19:37:40.626Z",
        "Description": "Example configuration revision description.",
        "Revision": 2
    }
}
```

Eliminazione di una configurazione MSK

Nella procedura seguente viene illustrato come eliminare una configurazione non collegata a un cluster. Non è possibile eliminare una configurazione collegata a un cluster.

1. Per eseguire questo esempio, sostituisci *configuration-arn* con l'ARN ottenuto al momento della creazione della configurazione. Se l'ARN non è stato salvato al momento della creazione della configurazione, è possibile utilizzare il comando `list-configurations` per elencare tutte le configurazioni presenti nell'account. La configurazione desiderata viene visualizzata nell'elenco di risposta. L'ARN della configurazione viene visualizzato anche in tale elenco.

```
aws kafka delete-configuration --arn configuration-arn
```

2. Di seguito è riportato un esempio di una risposta corretta dopo l'esecuzione di questo comando.

```
{
    "arn": " arn:aws:kafka:us-east-1:123456789012:configuration/SomeTest/
abcdabcd-1234-abcd-1234-abcd123e8e8e-1",
    "state": "DELETING"
}
```

Per descrivere una configurazione MSK

1. Il comando seguente restituisce i metadati relativi alla configurazione. Per ottenere una descrizione dettagliata della configurazione, eseguire `describe-configuration-revision`.

   Per eseguire questo esempio, sostituisci *`configuration-arn`* con l'ARN ottenuto al momento della creazione della configurazione. Se l'ARN non è stato salvato al momento della creazione della configurazione, è possibile utilizzare il comando `list-configurations` per elencare tutte le configurazioni presenti nell'account. La configurazione desiderata viene visualizzata nell'elenco di risposta. L'ARN della configurazione viene visualizzato anche in tale elenco.

   ```
   aws kafka describe-configuration --arn configuration-arn
   ```

2. Di seguito è riportato un esempio di una risposta corretta dopo l'esecuzione di questo comando.

   ```
   {
       "Arn": "arn:aws:kafka:us-east-1:123456789012:configuration/SomeTest/abcdabcd-
   abcd-1234-abcd-abcd123e8e8e-1",
       "CreationTime": "2019-05-21T00:54:23.591Z",
       "Description": "Example configuration description.",
       "KafkaVersions": [
           "1.1.1"
       ],
       "LatestRevision": {
           "CreationTime": "2019-05-21T00:54:23.591Z",
           "Description": "Example configuration description.",
           "Revision": 1
       },
       "Name": "SomeTest"
   }
   ```

Per descrivere una revisione della configurazione MSK

Se utilizzi il comando `describe-configuration` per descrivere una configurazione MSK, visualizzerai i metadati della configurazione. Per ottenere una descrizione dettagliata della configurazione, utilizza il comando `describe-configuration-revision`.

- Esegui il comando seguente, sostituendo *`configuration-arn`* con l'ARN ottenuto al momento della creazione della configurazione. Se l'ARN non è stato salvato al momento della creazione

della configurazione, è possibile utilizzare il comando `list-configurations` per elencare tutte le configurazioni presenti nell'account. La configurazione desiderata viene visualizzata nell'elenco di risposta. L'ARN della configurazione viene visualizzato anche in tale elenco.

```
aws kafka describe-configuration-revision --arn configuration-arn --revision 1
```

Di seguito è riportato un esempio di una risposta corretta dopo l'esecuzione di questo comando.

```
{
    "Arn": "arn:aws:kafka:us-east-1:123456789012:configuration/SomeTest/abcdabcd-
abcd-1234-abcd-abcd123e8e8e-1",
    "CreationTime": "2019-05-21T00:54:23.591Z",
    "Description": "Example configuration description.",
    "Revision": 1,
    "ServerProperties":
 "YXV0by5jcmVhdGUudG9waWNzLmVuYWJsZSA9IHRydWUKCgp6b29rZWVwZXIuY29ubmVjdGlvbi50aW1lb3V0Lm1zI
}
```

Il valore di `ServerProperties` è codificato con base64. Se si utilizza un decodificatore base64 (ad esempio, https://www.base64decode.org/) per decodificarlo manualmente, si ottiene il contenuto del file di configurazione originale utilizzato per creare la configurazione personalizzata. In questo caso, si ottiene quanto segue:

```
auto.create.topics.enable = true

log.roll.ms = 604800000
```

Per elencare tutte le configurazioni MSK nell'account per la regione corrente

- Esegui il comando seguente.

```
aws kafka list-configurations
```

Di seguito è riportato un esempio di una risposta corretta dopo l'esecuzione di questo comando.

```
{
    "Configurations": [
        {
```

```
            "Arn": "arn:aws:kafka:us-east-1:123456789012:configuration/SomeTest/
abcdabcd-abcd-1234-abcd-abcd123e8e8e-1",
            "CreationTime": "2019-05-21T00:54:23.591Z",
            "Description": "Example configuration description.",
            "KafkaVersions": [
                "1.1.1"
            ],
            "LatestRevision": {
                "CreationTime": "2019-05-21T00:54:23.591Z",
                "Description": "Example configuration description.",
                "Revision": 1
            },
            "Name": "SomeTest"
        },
        {
            "Arn": "arn:aws:kafka:us-east-1:123456789012:configuration/SomeTest/
abcdabcd-1234-abcd-1234-abcd123e8e8e-1",
            "CreationTime": "2019-05-03T23:08:29.446Z",
            "Description": "Example configuration description.",
            "KafkaVersions": [
                "1.1.1"
            ],
            "LatestRevision": {
                "CreationTime": "2019-05-03T23:08:29.446Z",
                "Description": "Example configuration description.",
                "Revision": 1
            },
            "Name": "ExampleConfigurationName"
        }
    ]
}
```

MSK Serverless

> **Note**
>
> MSK Serverless è disponibile nelle regioni Stati Uniti orientali (Ohio), Stati Uniti orientali (Virginia settentrionale), Stati Uniti occidentali (Oregon), Canada (Centrale), Asia Pacifico (Mumbai), Asia Pacifico (Singapore), Asia Pacifico (Sydney), Asia Pacifico (Tokyo), Asia Pacifico (Seoul), Europa (Francoforte), Europa (Stoccolma), Europa (Irlanda), Europa (Parigi) ed Europa (Londra).

MSK Serverless è un tipo di cluster per Amazon MSK che consente di eseguire Apache Kafka senza dover gestire e dimensionare la capacità del cluster. Fornisce e dimensiona automaticamente la capacità durante la gestione delle partizioni dell'argomento, in modo da poter trasmettere i dati senza pensare all'adeguamento o al dimensionamento dei cluster. MSK Serverless offre un modello tariffario basato sulla velocità di trasmissione effettiva, perciò ti viene addebitato soltanto l'utilizzo effettivo. Se le tue applicazioni richiedono una capacità di streaming on demand con aumento e riduzione automatiche, prendi in considerazione l'utilizzo di un cluster serverless.

MSK Serverless è completamente compatibile con Apache Kafka, quindi è possibile utilizzare qualsiasi applicazione client compatibile per produrre e utilizzare dati. Inoltre, si integra con i seguenti servizi:

- AWS PrivateLink per fornire connettività privata
- AWS Identity and Access Management (IAM) per l'autenticazione e l'autorizzazione utilizzando linguaggi Java e non Java. Per istruzioni sulla configurazione dei client per IAM, consulta Configurazione dei client per il Controllo degli accessi IAM.
- AWS Glue Registro degli schemi per la gestione degli schemi
- Servizio gestito da Amazon per Apache Flink per l'elaborazione di flussi basata su Apache Flink
- AWS Lambda per l'elaborazione degli eventi

> **Note**
>
> MSK Serverless richiede il Controllo degli accessi IAM per tutti i cluster. Le liste di controllo degli accessi (ACL) di Apache Kafka non sono supportate. Per ulteriori informazioni, consulta the section called “Controllo degli accessi IAM”.

Per informazioni sulle quote di servizio applicabili a MSK Serverless, consulta la sezione the section called “Quota per i cluster serverless”.

Per iniziare a utilizzare i cluster serverless e per ulteriori informazioni sulle opzioni di configurazione e monitoraggio per i cluster serverless, consulta le seguenti risorse.

Argomenti

- Guida introduttiva all'utilizzo dei cluster MSK Serverless
- Configurazione per cluster serverless
- Monitoraggio dei cluster serverless

Guida introduttiva all'utilizzo dei cluster MSK Serverless

Questo tutorial mostra un esempio di come creare un cluster MSK Serverless, creare un computer client in grado di accedervi e utilizzare il client per creare argomenti sul cluster e scrivere dati su tali argomenti. Questo esempio non rappresenta tutte le opzioni che è possibile scegliere quando si crea un cluster serverless. In diverse parti di questo esercizio verranno scelte opzioni predefinite per semplicità. Ciò non significa che siano le uniche opzioni che funzionano per la configurazione del cluster serverless. Puoi anche utilizzare l'API AWS CLI o Amazon MSK. Per ulteriori informazioni, consulta la documentazione di riferimento all'API di Amazon MSK 2.0.

Argomenti

- Passaggio 1: creazione di un cluster MSK Serverless
- Fase 2: creazione di un ruolo IAM
- Passaggio 3: creazione di un computer client
- Passaggio 4: creazione di un argomento di Apache Kafka
- Passaggio 5: produzione e utilizzo di dati
- Passaggio 6: eliminazione delle risorse

Passaggio 1: creazione di un cluster MSK Serverless

In questo passaggio, eseguirai due attività. Innanzitutto, si crea un cluster MSK Serverless con le impostazioni predefinite. In secondo luogo, si raccolgono informazioni sul cluster. Si tratta di

informazioni che ti occorreranno nei passaggi successivi, quando creerai un client in grado di inviare dati al cluster.

Creazione di un cluster serverless

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home.
2. Scegli Create cluster (Crea cluster).
3. Per Metodo di creazione, lascia selezionata l'opzione Creazione rapida. L'opzione Creazione rapida consente di creare un cluster serverless con le impostazioni predefinite.
4. In Nome del cluster, inserisci un nome descrittivo, ad esempio **`msk-serverless-tutorial-cluster`**.
5. In Proprietà generali del cluster, scegli Serverless come Tipo di cluster. Utilizza i valori predefiniti per le Proprietà generali del cluster rimanenti.
6. Nota la tabella in Tutte le impostazioni del cluster. Questa tabella elenca i valori predefiniti per impostazioni importanti come rete e disponibilità e indica se è possibile modificare ogni impostazione dopo aver creato il cluster. Per modificare un'impostazione prima di creare il cluster, è necessario scegliere l'opzione Creazione personalizzata in Metodo di creazione.

   > **Note**
   >
   > Ai cluster MSK Serverless è possibile connettere client da un massimo di cinque VPC diversi. Per aiutare le applicazioni client a passare a un'altra zona di disponibilità in caso di interruzione, è necessario specificare almeno due sottoreti in ogni VPC.

7. Scegli Create cluster (Crea cluster).

Raccolta delle informazioni sul cluster

1. Nella pagina Riepilogo del cluster, scegli Visualizza informazioni sul client. Questo pulsante rimane disattivato fino al termine della creazione del cluster da parte di Amazon MSK. Potrebbe essere necessario attendere qualche minuto prima che il pulsante diventi attivo e possa essere selezionato.
2. Copia la stringa sotto l'etichetta Endpoint. Questa è la stringa del tuo server di bootstrap.
3. Scegliere la scheda Properties (Proprietà).

4. Nella sezione Impostazioni di rete, copia gli ID delle sottoreti e del gruppo di sicurezza e salvali perché queste informazioni ti serviranno in seguito per creare un computer client.
5. Scegli una delle sottoreti. Si apre la console Amazon VPC. Cerca l'ID dell'Amazon VPC associato al VPC della sottorete. Salva questo ID dell'Amazon VPC per utilizzarlo in futuro.

Fase successiva

Fase 2: creazione di un ruolo IAM

Fase 2: creazione di un ruolo IAM

In questo passaggio, eseguirai due attività. La prima attività consiste nel creare una policy IAM che consenta l'accesso alla creazione di argomenti nel cluster e all'invio di dati a tali argomenti. La seconda attività consiste nel creare un ruolo IAM e associarvi questa policy. In un passaggio successivo, si crea un computer client che assume questo ruolo e lo utilizza per creare un argomento nel cluster e per inviare dati a quell'argomento.

Creazione di una policy IAM che consenta di creare argomenti e scrivere su di essi

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro di navigazione, seleziona Policy.
3. Scegliere Create Policy (Crea policy).
4. Scegli la scheda JSON, quindi sostituisci il JSON nella finestra dell'editor con il seguente JSON.

   Sostituisci *region* con il codice della Regione AWS in cui hai creato il cluster. Sostituisci *Account-ID* con il tuo ID account. *msk-serverless-tutorial-cluster*Sostituiscilo con il nome del tuo cluster serverless.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:Connect",
                "kafka-cluster:AlterCluster",
                "kafka-cluster:DescribeCluster"
            ],
            "Resource": [
```

```
                "arn:aws:kafka:region:Account-ID:cluster/msk-serverless-tutorial-
cluster/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:*Topic*",
                "kafka-cluster:WriteData",
                "kafka-cluster:ReadData"
            ],
            "Resource": [
                "arn:aws:kafka:region:Account-ID:topic/msk-serverless-tutorial-
cluster/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:AlterGroup",
                "kafka-cluster:DescribeGroup"
            ],
            "Resource": [
                "arn:aws:kafka:region:Account-ID:group/msk-serverless-tutorial-
cluster/*"
            ]
        }
    ]
}
```

Per ricevere istruzioni su come scrivere policy sicure, consulta la pagina the section called "Controllo degli accessi IAM".

5. Scegliere Next: Tags (Successivo: Tag).
6. Scegliere Next:Review (Successivo: Rivedi).
7. Per il nome della policy, inserisci un nome descrittivo, ad esempio **`msk-serverless-tutorial-policy`**.
8. Scegli Crea policy.

Creazione di un ruolo IAM e collegamento della policy al ruolo

1. Nel riquadro di navigazione, seleziona Ruoli.
2. Scegli Crea ruolo.
3. In Casi di utilizzo comuni, scegli EC2, quindi scegli Successivo: autorizzazioni.
4. Nella casella di ricerca, inserisci il nome della policy creata in precedenza per questo tutorial. Seleziona quindi la casella a sinistra della policy.
5. Scegliere Next: Tags (Successivo: Tag).
6. Scegliere Next:Review (Successivo: Rivedi).
7. Per il nome del ruolo, inserisci un nome descrittivo, ad esempio **`msk-serverless-tutorial-role`**.
8. Scegli Crea ruolo.

Fase successiva

Passaggio 3: creazione di un computer client

Passaggio 3: creazione di un computer client

In questo passaggio, eseguirai due attività. La prima operazione consiste nel creare un'istanza Amazon EC2 da utilizzare come computer client Apache Kafka. La seconda attività consiste nell'installare gli strumenti Java e Apache Kafka sul computer.

Per creare un computer client

1. Aprire la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Scegliere Launch Instance (Avvia istanza).
3. Inserisci un Nome descrittivo per il computer client, ad esempio **`msk-serverless-tutorial-client`**.
4. Lascia Amazon Linux 2 AMI (HVM) - Kernel 5.10, tipo di volume SSD selezionato per Tipo di Amazon Machine Image (AMI).
5. Lascia selezionato il tipo di istanza t2.micro.
6. In Coppia di chiavi (accesso), scegli Crea una nuova coppia di chiavi. Inserisci **`MSKServerlessKeyPair`** per Nome coppia di chiavi. Quindi scegli Scarica coppia di chiavi. In alternativa, è possibile utilizzare una coppia di chiavi esistente.

7. Per Impostazioni di rete, scegli Modifica.
8. In VPC, inserisci l'ID del cloud privato virtuale (VPC) per il cluster serverless. Si tratta del VPC basato sul servizio Amazon VPC il cui ID è stato salvato dopo la creazione del cluster.
9. Per Sottorete, scegli la sottorete di cui hai salvato l'ID dopo aver creato il cluster.
10. Per Firewall (gruppi di sicurezza), seleziona il gruppo di sicurezza associato al cluster. Questo valore funziona se il gruppo di sicurezza ha una regola in entrata che consente il traffico dal gruppo di sicurezza verso sé stesso. Con questa regola, i membri dello stesso gruppo di sicurezza possono comunicare tra loro. Per ulteriori informazioni, consulta la pagina Security group rules nella Guida per gli sviluppatori di Amazon VPC.
11. Espandi la sezione Dettagli avanzati e scegli il ruolo IAM che hai creato nel Fase 2: creazione di un ruolo IAM.
12. Scegli Avvia.
13. Nel riquadro di navigazione a sinistra, scegliere Instances (Istanze). Quindi scegli la casella di controllo nella riga che rappresenta l'istanza Amazon EC2 appena creata. D'ora in avanti, chiameremo questa istanza computer client.
14. Scegli Connetti e segui le istruzioni per connetterti al computer client.

Configurazione degli strumenti client Apache Kafka sul computer client

1. Per installare Java, esegui il comando seguente sul computer client:

```
sudo yum -y install java-11
```

2. Per recuperare gli strumenti di Apache Kafka necessari per creare argomenti e inviare dati, esegui i seguenti comandi:

```
wget https://archive.apache.org/dist/kafka/2.8.1/kafka_2.12-2.8.1.tgz
```

```
tar -xzf kafka_2.12-2.8.1.tgz
```

3. Vai alla directory `kafka_2.12-2.8.1/libs`, quindi esegui il comando per scaricare il file JAR IAM di Amazon MSK. Il file JAR IAM di Amazon MSK consente al computer client di accedere al cluster.

```
wget https://github.com/aws/aws-msk-iam-auth/releases/download/v1.1.1/aws-msk-iam-auth-1.1.1-all.jar
```

4. Vai alla directory `kafka_2.12-2.8.1/bin`. Copia le impostazioni delle proprietà seguenti e incollale in un nuovo file. Assegna al file il nome `client.properties` e salvalo.

```
security.protocol=SASL_SSL
sasl.mechanism=AWS_MSK_IAM
sasl.jaas.config=software.amazon.msk.auth.iam.IAMLoginModule required;
sasl.client.callback.handler.class=software.amazon.msk.auth.iam.IAMClientCallbackHandler
```

Fase successiva

Passaggio 4: creazione di un argomento di Apache Kafka

Passaggio 4: creazione di un argomento di Apache Kafka

In questo passaggio, si utilizza il computer client creato in precedenza per creare un argomento sul cluster serverless.

Creazione di un argomento e scrittura di dati su di esso

1. Nel comando `export` seguente, sostituisci *`my-endpoint`* con la stringa del server di bootstrap che hai salvato dopo aver creato il cluster. Quindi, vai alla directory `kafka_2.12-2.8.1/bin` sul computer client ed esegui il comando `export`.

```
export BS=my-endpoint
```

2. Esegui il comando seguente per creare un argomento chiamato `msk-serverless-tutorial`.

```
<path-to-your-kafka-installation>/bin/kafka-topics.sh --bootstrap-server $BS
 --command-config client.properties --create --topic msk-serverless-tutorial --
partitions 6
```

Fase successiva

Passaggio 5: produzione e utilizzo di dati

Passaggio 5: produzione e utilizzo di dati

In questo passaggio, si producono e si utilizzano dati utilizzando l'argomento creato nel passaggio precedente.

Per produrre e consumare messaggi

1. Esegui il comando seguente per creare un produttore della console.

```
<path-to-your-kafka-installation>/bin/kafka-console-producer.sh --broker-list $BS
 --producer.config client.properties --topic msk-serverless-tutorial
```

2. Immettere qualsiasi messaggio desiderato e premere Enter (Invio). Ripetere questa fase due o tre volte. Ogni volta che immetti una riga e premi Invio, tale riga viene inviata al cluster Apache Kafka come un messaggio separato.
3. Mantenere aperta la connessione al computer client, quindi aprire una seconda connessione separata al computer in una nuova finestra.
4. Utilizza la tua seconda connessione al computer client per creare un utente della console eseguendo il comando seguente. Sostituisci *my-endpoint* con la stringa del server di bootstrap che hai salvato dopo aver creato il cluster.

```
<path-to-your-kafka-installation>/bin/kafka-console-consumer.sh --bootstrap-
server my-endpoint --consumer.config client.properties --topic msk-serverless-
tutorial --from-beginning
```

   Si iniziano a vedere i messaggi immessi in precedenza quando è stato utilizzato il comando produttore della console.
5. Immettere altri messaggi nella finestra del produttore e guardali apparire nella finestra del consumatore.

Fase successiva

Passaggio 6: eliminazione delle risorse

Passaggio 6: eliminazione delle risorse

In questo passaggio, elimini le risorse che hai creato in questo tutorial.

Eliminazione del cluster

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home.
2. Nell'elenco dei cluster, scegli il cluster che hai creato per questo tutorial.
3. In Operazioni, scegli Elimina cluster.

4. Inserisci `delete` nel campo e scegli Elimina.

Arresto del computer client

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nell'elenco delle istanze Amazon EC2, scegli il computer client creato per questo tutorial.
3. Scegli Stato istanza, quindi scegli Termina istanza.
4. Scegliere Terminate (Termina).

Eliminazione del ruolo e della policy IAM

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro di navigazione, seleziona Ruoli.
3. Nella casella di ricerca, inserisci il nome del ruolo IAM creato per questo tutorial.
4. Seleziona il ruolo. Quindi scegli Elimina ruolo e conferma l'eliminazione.
5. Nel riquadro di navigazione, seleziona Policy.
6. Nella casella di ricerca, inserisci il nome della policy creata per questo tutorial.
7. Scegli la policy per aprirne la pagina di riepilogo. Nella pagina di Riepilogo della policy, seleziona Elimina policy.
8. Scegli Delete (Elimina).

Configurazione per cluster serverless

Amazon MSK imposta le proprietà di configurazione del broker per i cluster serverless. Non è possibile modificare queste impostazioni delle proprietà di configurazione del broker. Tuttavia, è possibile impostare le proprietà di configurazione degli argomenti seguenti.

Proprietà di configurazione	Predefinita	Modificabile	Valore massimo consentito
cleanup.policy	Eliminazione	Sì, ma solo al momento della creazione dell'argomento	

Proprietà di configurazione	Predefinita	Modificabile	Valore massimo consentito
compression.type	Producer	Sì	
max.message.bytes	1048588	Sì	8 MiB
message.timestamp.difference.max.ms	long.max	Sì	
message.timestamp.type	CreateTime	Sì	
retention.bytes	250 GiB	Sì	250 GiB
retention.ms	7 giorni	Sì	Illimitato

Per impostare o modificare le proprietà di configurazione a livello dell'argomento per argomenti nuovi o esistenti, puoi utilizzare anche i comandi di Apache Kafka. Per ulteriori informazioni sulle proprietà di configurazione a livello di argomento ed esempi di come impostarle, consulta la pagina Topic-Level Configs nella documentazione ufficiale di Apache Kafka.

Monitoraggio dei cluster serverless

Amazon MSK si integra con Amazon per CloudWatch consentirti di raccogliere, visualizzare e analizzare i parametri per il tuo cluster MSK Serverless. I parametri mostrati nella tabella seguente sono disponibili per tutti i cluster serverless. Poiché questi parametri sono pubblicati come punti di dati individuali per ogni partizione dell'argomento, consigliamo di visualizzarle come statistiche "SUM" per ottenere una visualizzazione a livello di argomento.

Amazon MSK pubblica i `PerSec` parametri con una frequenza di una volta CloudWatch al minuto. Ciò significa che la statistica "SUM" per un periodo di un minuto rappresenta accuratamente i dati al secondo per i parametri `PerSec`. Per raccogliere dati al secondo per un periodo superiore a un minuto, usa la seguente espressione matematica:. CloudWatch `m1 * 60/PERIOD(m1)`

Parametri disponibili al livello di monitoraggio DEFAULT

Nome	Quando visibile	Dimensioni	Descrizione
`BytesInPerSec`	Dopo che un produttore ha scritto su un argomento	Nome del cluster, argomento	Il numero di byte al secondo ricevuti dai client. Questo parametro è disponibile per ogni argomento.
`BytesOutPerSec`	Dopo che un gruppo di consumatori ha utilizzato un argomento	Nome del cluster, argomento	Il numero di byte al secondo inviati ai client. Questo parametro è disponibi le per ogni argomento.
`FetchMess ageConver sionsPerSec`	Dopo che un gruppo di consumatori ha utilizzato un argomento	Nome del cluster, argomento	Il numero di conversioni dei messaggi di recupero al secondo per l'argomen to.
`Estimated MaxTimeLag`	Dopo che un gruppo di consumatori ha utilizzato un argomento	Nome del cluster, gruppo di consumatori, argomento	Una stima temporale della metrica. MaxOffsetLag
`MaxOffsetLag`	Dopo che un gruppo di consumatori ha utilizzato un argomento	Nome del cluster, gruppo di consumatori, argomento	Il ritardo massimo di offset su tutte le partizioni di un argomento.
`MessagesI nPerSec`	Dopo che un produttore ha scritto su un argomento	Nome del cluster, argomento	Il numero di messaggi in entrata al secondo per l'argomento.

Nome	Quando visibile	Dimensioni	Descrizione
`ProduceMessageConversionsPerSec`	Dopo che un produttore ha scritto su un argomento	Nome del cluster, argomento	Il numero di conversioni di messaggi di produzione al secondo per l'argomento.
`SumOffsetLag`	Dopo che un gruppo di consumatori ha utilizzato un argomento	Nome del cluster, gruppo di consumatori, argomento	Il ritardo di offset aggregato per tutte le partizioni di un argomento.

Visualizzazione dei parametri di MSK Serverless

1. Accedi AWS Management Console e apri la CloudWatch console all'indirizzo https://console.aws.amazon.com/cloudwatch/.
2. Nel riquadro di navigazione, in Parametri, scegli Tutti i parametri.
3. Nei parametri, cerca il termine **`kafka`**.
4. Per visualizzare parametri diversi, scegli AWS/Kafka / Nome del cluster, argomento oppure AWS/Kafka / Nome del cluster, gruppo di consumatori, argomento.

MSK Connect

Cos'è MSK Connect?

MSK Connect è una funzionalità di Amazon MSK che semplifica lo streaming di dati da e verso i cluster Apache Kafka a vantaggio degli sviluppatori. MSK Connect utilizza Kafka Connect 2.7.1, un framework open source, per connettere i cluster Apache Kafka con sistemi esterni come database, indici di ricerca e file system. Con MSK Connect, è possibile implementare connettori completamente gestiti creati per Kafka Connect che trasferiscono o estraggono dati da archivi di dati popolari come Amazon S3 e Amazon Service. OpenSearch È possibile implementare connettori sviluppati da terze parti come Debezium per eseguire lo streaming dei log delle modifiche dai database a un cluster Apache Kafka, oppure implementare un connettore esistente senza modifiche al codice. I connettori si dimensionano automaticamente in base alle variazioni di carico e ti verranno addebitate soltanto le risorse che effettivamente utilizzi.

Utilizza i connettori di origine per importare dati da sistemi esterni nei tuoi argomenti. Con i connettori sink, è possibile esportare i dati dai propri argomenti a sistemi esterni.

MSK Connect supporta connettori per qualsiasi cluster Apache Kafka con connettività a un Amazon VPC, che si tratti di un cluster MSK o di un cluster Apache Kafka ospitato in modo indipendente.

MSK Connect monitora continuamente l'integrità e lo stato di consegna dei connettori, corregge e gestisce l'hardware sottostante e dimensiona automaticamente i connettori in base alle variazioni della velocità di trasmissione effettiva.

Per le nozioni di base su MSK Connect, consulta la pagina the section called “Nozioni di base”.

Per ulteriori informazioni sulle AWS risorse che è possibile creare con MSK Connect, vedere the section called “Connectors (Connettori)”the section called “Plug-in”, ethe section called “Worker”.

Per informazioni sull'API di MSK Connect, consulta la documentazione di riferimento sull'API di Amazon MSK Connect.

Guida introduttiva all'utilizzo di MSK Connect

Questo è un step-by-step tutorial che utilizza AWS Management Console per creare un cluster MSK e un connettore sink che invia i dati dal cluster a un bucket S3.

Argomenti

Passaggio 1: configurazione delle risorse

In questo passaggio crei le seguenti risorse necessarie per questo scenario introduttivo:

- Un bucket S3 che funge da destinazione per la ricezione dei dati dal connettore.
- Un cluster MSK a cui inviare i dati. Il connettore leggerà i dati da questo cluster e li invierà al bucket S3 di destinazione.
- Un ruolo IAM che consente al connettore di scrivere nel bucket S3 di destinazione.
- Un endpoint Amazon VPC per consentire l'invio di dati dall'Amazon VPC che include il cluster e il connettore ad Amazon S3.

Creazione del bucket S3

1. Accedi AWS Management Console e apri la console Amazon S3 all'indirizzo https://console.aws.amazon.com/s3/.
2. Seleziona Crea bucket.
3. Per il nome del bucket, specifica un nome descrittivo, ad esempio `mkc-tutorial-destination-bucket`.
4. Scorri verso il basso e scegli Crea bucket.
5. Nell'elenco dei bucket, scegli il bucket appena creato.
6. Scegliere Create folder (Crea cartella).
7. Inserisci `tutorial` come nome della cartella, quindi scorri verso il basso e scegli Crea cartella.

Creazione del cluster

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Nel riquadro a sinistra, in Cluster MSK, scegli Cluster.

3. Scegli Create cluster (Crea cluster).
4. Scegli Creazione personalizzata.
5. Come nome del cluster, specifica `mkc-tutorial-cluster`.
6. In Proprietà generali del cluster, scegli Assegnato come tipo di cluster.
7. In Rete, scegli un Amazon VPC. Seleziona quindi le zone di disponibilità e le sottoreti che desideri utilizzare. Ricorda gli ID dell'Amazon VPC e delle sottoreti che hai selezionato perché ne avrai bisogno più avanti in questo tutorial.
8. In Metodi di controllo degli accessi, assicurati che sia selezionato soltanto Accesso non autenticato.
9. In Crittografia, assicurati che sia selezionato solo Non crittografato.
10. Continua con la procedura guidata, quindi scegli Crea cluster. In questo modo si accede alla pagina Dettagli per il cluster. In quella pagina, in Gruppi di sicurezza applicati, trova l'ID del gruppo di sicurezza. Ricorda quell'ID perché ne avrai bisogno più avanti in questo tutorial.

Creazione del ruolo IAM che può scrivere nel bucket di destinazione

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro a sinistra, in Gestione degli accessi, scegli Ruoli.
3. Scegli Crea ruolo.
4. In Oppure seleziona un servizio per visualizzarne i casi d'uso, scegli S3.
5. Scorri verso il basso e in Seleziona il tuo caso d'uso scegli nuovamente S3.
6. Scegli Next: Permissions (Successivo: Autorizzazioni).
7. Scegli Crea policy. Nel browser si apre una nuova scheda nella quale potrai creare la policy. Lascia aperta la scheda originale per la creazione del ruolo perché vi torneremo più tardi.
8. Scegli la scheda JSON, quindi sostituisci il testo nella finestra con la policy seguente.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListAllMyBuckets"
      ],
      "Resource": "arn:aws:s3:::*"
```

```
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket",
        "s3:GetBucketLocation",
        "s3:DeleteObject"
      ],
      "Resource": "arn:aws:s3:::<my-tutorial-destination-bucket>"
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:AbortMultipartUpload",
        "s3:ListMultipartUploadParts",
        "s3:ListBucketMultipartUploads"
      ],
      "Resource": "*"
    }
  ]
}
```

9. Scegliere Next: Tags (Successivo: Tag).
10. Scegliere Next:Review (Successivo: Rivedi).
11. Inserisci `mkc-tutorial-policy` come nome della policy, quindi scorri verso il basso e scegli Crea policy.
12. Torna alla scheda del browser in cui stavi creando il ruolo e scegli il pulsante di aggiornamento.
13. Trova la policy `mkc-tutorial-policy` e selezionala facendo clic sul pulsante alla sua sinistra.
14. Scegliere Next: Tags (Successivo: Tag).
15. Scegliere Next:Review (Successivo: Rivedi).
16. Inserisci `mkc-tutorial-role` come nome del ruolo ed elimina il testo nella casella della descrizione.
17. Scegli Crea ruolo.

Autorizzazione di MSK Connect ad assumere il ruolo

1. Nella console IAM, nel riquadro a sinistra, in Gestione degli accessi, scegli Ruoli.

2. Trova il ruolo `mkc-tutorial-role` e selezionalo.
3. Nel Riepilogo del ruolo, scegli la scheda Relazioni di attendibilità.
4. Seleziona Modifica relazione di attendibilità.
5. Sostituisci la policy di attendibilità esistente con il JSON seguente.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "kafkaconnect.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

6. Scegli Update Trust Policy (Aggiorna policy di trust).

Creazione di un endpoint VPC dal VPC del cluster ad Amazon S3

1. Apri alla console Amazon VPC all'indirizzo https://console.aws.amazon.com/vpc/.
2. Nel riquadro a sinistra, scegli Endpoint.
3. Seleziona Crea endpoint.
4. In Nome del servizio, scegli il servizio com.amazonaws.us-east-1.s3 e il tipo di Gateway.
5. Scegli il VPC del cluster, quindi seleziona la casella a sinistra della tabella di routing associata alle sottoreti del cluster.
6. Seleziona Crea endpoint.

Fase successiva

Passaggio 2: creazione di un plug-in personalizzato

Passaggio 2: creazione di un plug-in personalizzato

Un plug-in contiene il codice che definisce la logica del connettore. In questo passaggio crei un plug-in personalizzato con il codice per il connettore sink Amazon S3 Lenses. In un passaggio

successivo, quando creerai il connettore MSK, specificherai che il relativo codice si trova in questo plug-in personalizzato. È possibile utilizzare lo stesso plug-in per creare più connettori MSK con configurazioni diverse.

Creazione del plug-in personalizzato

1. Scarica il connettore S3.
2. Carica il file ZIP in un bucket S3 al quale puoi accedere. Per istruzioni su come caricare i file in Amazon S3, consulta la pagina Uploading objects nella Guida per l'utente di Amazon S3.
3. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
4. Nel riquadro a sinistra, espandi MSK Connect, quindi scegli Plug-in personalizzati.
5. Scegli Crea plug-in personalizzato.
6. Seleziona Sfoglia S3.
7. Nell'elenco dei bucket, trova il bucket in cui hai caricato il file ZIP e selezionalo.
8. Nell'elenco degli oggetti nel bucket, seleziona il pulsante di opzione a sinistra del file ZIP, quindi seleziona il pulsante con l'etichetta Scegli.
9. Inserisci `mkc-tutorial-plugin` come nome del plug-in personalizzato, quindi scegli Crea plug-in personalizzato.

Potrebbero essere necessari AWS alcuni minuti per completare la creazione del plug-in personalizzato. Al termine del processo di creazione, nella parte superiore della finestra del browser viene visualizzato il seguente messaggio in un banner.

```
Custom plugin mkc-tutorial-plugin was successfully created
The custom plugin was created. You can now create a connector using this custom plugin.
```

Fase successiva

Passaggio 3: creazione del computer client e dell'argomento Apache Kafka

Passaggio 3: creazione del computer client e dell'argomento Apache Kafka

In questo passaggio viene creata un'istanza Amazon EC2 da utilizzare come istanza client Apache Kafka. Quindi usi questa istanza per creare un argomento sul cluster.

Per creare un computer client

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Scegliere Launch Instances (Avvia istanze).
3. Inserisci un Nome per il computer client, ad esempio **`mkc-tutorial-client`**.
4. Lascia Amazon Linux 2 AMI (HVM) - Kernel 5.10, tipo di volume SSD selezionato per Tipo di Amazon Machine Image (AMI).
5. Scegli il tipo di istanza t2.xlarge.
6. In Coppia di chiavi (accesso), scegli Crea una nuova coppia di chiavi. Inserisci **`mkc-tutorial-key-pair`** in Nome della coppia di chiavi, quindi scegli Scarica coppia di chiavi. In alternativa, è possibile utilizzare una coppia di chiavi esistente.
7. Scegliere Launch Instance (Avvia istanza).
8. Scegliere View Instances (Vedi istanze). Quindi, nella colonna Gruppi di sicurezza, scegli il gruppo di sicurezza associato alla nuova istanza. Copia l'ID del gruppo di sicurezza e salvalo per un secondo momento.

Autorizzazione del client appena creato all'invio di dati al cluster

1. Apri alla console Amazon VPC all'indirizzo https://console.aws.amazon.com/vpc/.
2. Nel riquadro a sinistra, in Sicurezza, scegli Gruppi di sicurezza. Nella colonna ID del gruppo di sicurezza, trova il gruppo di sicurezza del cluster. Hai salvato l'ID di questo gruppo di sicurezza quando hai creato il cluster in the section called “Passaggio 1: configurazione delle risorse”. Scegli questo gruppo di sicurezza selezionando la casella a sinistra della riga. Assicurati che nessun altro gruppo di sicurezza sia selezionato contemporaneamente.
3. Nella sezione inferiore della pagina, scegli la scheda Regole in entrata.
4. Sceglere Edit inbound rules (Modifica regole in entrata).
5. In basso a sinistra dello schermo, scegli Aggiungi regola.
6. Nella nuova regola, scegliere All traffic (Tutto il traffico) nella colonna Type (Tipo) . Nel campo a destra della colonna Origine, inserisci l'ID del gruppo di sicurezza del computer client. Questo è l'ID del gruppo di sicurezza che hai salvato dopo aver creato il computer client.
7. Scegliere Salva regole. Il cluster MSK ora accetterà tutto il traffico proveniente dal client creato nella procedura precedente.

Per creare un argomento

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nella tabella delle istanze, scegli `mkc-tutorial-client`.
3. Nella parte superiore dello schermo, scegli Connetti, quindi segui le istruzioni per connetterti all'istanza.
4. Installa Java sull'istanza client eseguendo il seguente comando:

```
sudo yum install java-1.8.0
```

5. Eseguire il seguente comando per scaricare Apache Kafka.

```
wget https://archive.apache.org/dist/kafka/2.2.1/kafka_2.12-2.2.1.tgz
```

> **Note**
>
> Se desideri utilizzare un sito mirror diverso da quello utilizzato in questo comando, puoi sceglierne uno diverso sul sito Web di Apache .

6. Eseguire il comando seguente nella directory in cui è stato scaricato il file TAR nella fase precedente.

```
tar -xzf kafka_2.12-2.2.1.tgz
```

7. Passare alla directory kafka_2.12-2.2.1 .
8. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
9. Nel riquadro a sinistra, scegli Cluster, quindi scegli il nome `mkc-tutorial-cluster`.
10. Scegli Visualizza le informazioni sul client.
11. Copia la stringa di connessione Non crittografato.
12. Seleziona Fatto.
13. Eseguite il comando seguente sull'istanza del client (`mkc-tutorial-client`), sostituendolo *`bootstrapServerString`*con il valore salvato quando avete visualizzato le informazioni sul client del cluster.

```
<path-to-your-kafka-installation>/bin/kafka-topics.sh --create --bootstrap-
server bootstrapServerString --replication-factor 2 --partitions 1 --topic mkc-
tutorial-topic
```

Se il comando va a buon fine, viene visualizzato il seguente messaggio: `Created topic mkc-tutorial-topic.`

Fase successiva

Passaggio 4: creazione del connettore

Passaggio 4: creazione del connettore

Creazione del connettore

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Nel riquadro a sinistra, espandi MSK Connect, quindi scegli Connettori.
3. Scegli Create connector (Crea connettore).
4. Nell'elenco dei plugin, scegli `mkc-tutorial-plugin`, quindi scegli Avanti.
5. Per il nome del connettore, inserisci `mkc-tutorial-connector`.
6. Nell'elenco dei cluster, scegli `mkc-tutorial-cluster`.
7. Copia la configurazione seguente e incollala nel campo di configurazione del connettore.

```
connector.class=io.confluent.connect.s3.S3SinkConnector
s3.region=us-east-1
format.class=io.confluent.connect.s3.format.json.JsonFormat
flush.size=1
schema.compatibility=NONE
tasks.max=2
topics=mkc-tutorial-topic
partitioner.class=io.confluent.connect.storage.partitioner.DefaultPartitioner
storage.class=io.confluent.connect.s3.storage.S3Storage
s3.bucket.name=<my-tutorial-destination-bucket>
topics.dir=tutorial
```

8. In Autorizzazioni di accesso, scegli `mkc-tutorial-role`.

9. Seleziona Successivo. Nella pagina Sicurezza, scegli di nuovo Avanti.
10. Nella pagina Log, seleziona Avanti.
11. In Rivedi e crea, scegli Crea connettore.

Fase successiva

Passaggio 5: invio dei dati

Passaggio 5: invio dei dati

In questo passaggio si inviano i dati all'argomento Apache Kafka creato in precedenza, quindi si cercano gli stessi dati nel bucket S3 di destinazione.

Invio dei dati al cluster MSK

1. Nella cartella `bin` dell'installazione di Apache Kafka sull'istanza client, crea un file di testo denominato `client.properties` con i seguenti contenuti.

```
security.protocol=PLAINTEXT
```

2. Esegui il comando seguente per creare un produttore della console. Sostituisci *`BootstrapBrokerString`*con il valore ottenuto quando hai eseguito il comando precedente.

```
<path-to-your-kafka-installation>/bin/kafka-console-producer.sh --broker-
list BootstrapBrokerString --producer.config client.properties --topic mkc-
tutorial-topic
```

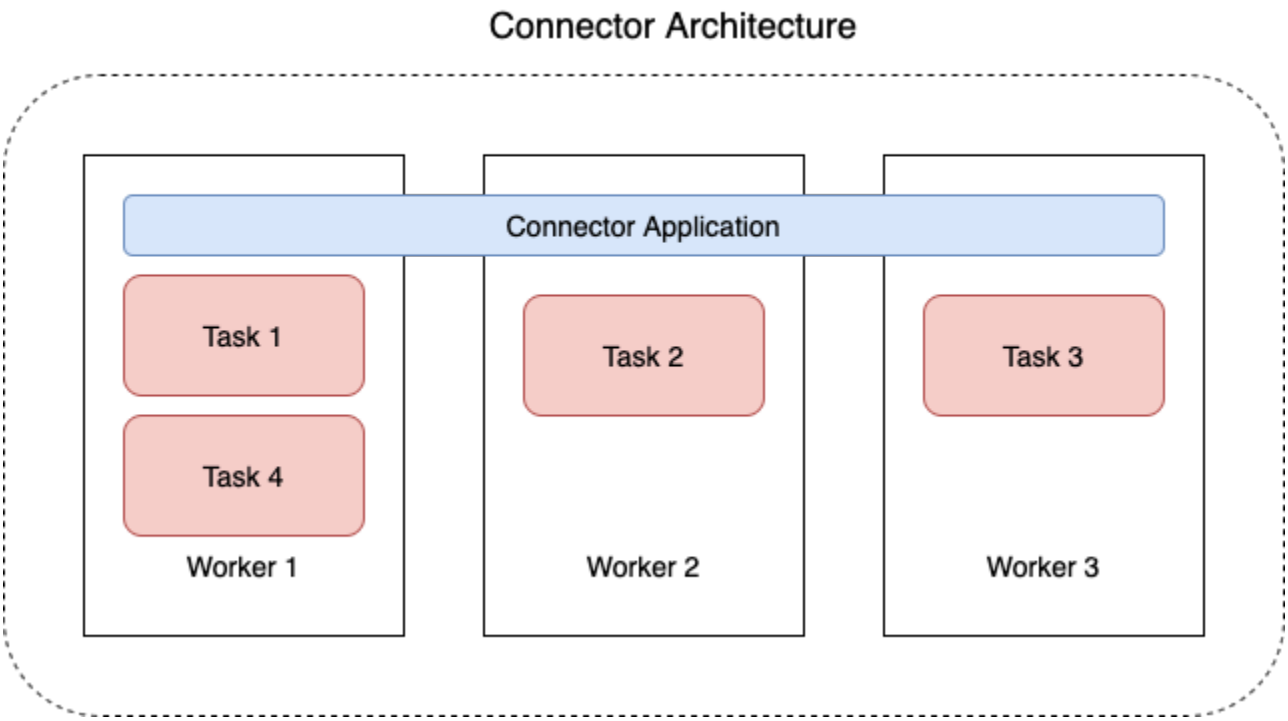
3. Immettere qualsiasi messaggio desiderato e premere Enter (Invio). Ripetere questa fase due o tre volte. Ogni volta che si immette una riga e si preme Enter (Invio), tale riga viene inviata al cluster Apache Kafka come un messaggio separato.
4. Cerca nel bucket Amazon S3 di destinazione per trovare i messaggi inviati nel passaggio precedente.

Connectors (Connettori)

Un connettore integra sistemi esterni e servizi Amazon con Apache Kafka copiando continuamente i dati in streaming da un'origine dati nel cluster Apache Kafka o dal cluster in un sink di dati. Un

connettore può anche eseguire operazioni logiche leggere come la trasformazione, la conversione del formato o il filtraggio dei dati prima di consegnarli a una destinazione. I connettori di origine estraggono i dati da un'origine dati e li inviano al cluster, mentre i connettori sink estraggono i dati dal cluster e li inviano a un sink di dati.

Nel diagramma seguente viene illustrata l'architettura di un connettore. Un worker è un processo di macchina virtuale Java (JVM) che esegue la logica del connettore. Ogni worker crea una serie di attività che vengono eseguite in thread paralleli e si occupano di copiare i dati. Le attività non memorizzano lo stato e possono quindi essere avviate, interrotte o riavviate in qualsiasi momento per fornire una pipeline di dati resiliente e scalabile.



Capacità del connettore

La capacità totale di un connettore dipende dal numero di worker del connettore e dal numero di MSK Connect Unit (MCU) per worker. Ogni MCU rappresenta 1 vCPU di elaborazione e 4 GiB di memoria. La memoria MCU riguarda la memoria totale di un'istanza worker e non la memoria heap in uso.

Gli operatori di MSK Connect utilizzano gli indirizzi IP nelle sottoreti fornite dal cliente. Ogni lavoratore utilizza un indirizzo IP da una delle sottoreti fornite dal cliente. È necessario assicurarsi di disporre di un numero sufficiente di indirizzi IP disponibili nelle sottoreti fornite a una CreateConnector richiesta per tenere conto della capacità specificata, specialmente quando si scalano automaticamente i connettori in cui il numero di lavoratori può variare.

Per creare un connettore, è necessario scegliere tra una delle due modalità di capacità seguenti.

- Assegnato: scegli questa modalità se conosci i requisiti di capacità del tuo connettore. Specifica due valori:
  - Il numero di worker.
  - Il numero di MCU per worker.
- Dimensionamento automatico: scegli questa modalità se i requisiti di capacità del connettore sono variabili o se non li conosci in anticipo. Quando si utilizza la modalità con dimensionamento automatico, Amazon MSK Connect sovrascrive la proprietà `tasks.max` del connettore con un valore proporzionale al numero di worker in esecuzione nel connettore e al numero di MCU per worker.

  Devi specificare tre set di valori:
  - Il numero minimo e massimo di worker.
  - Le percentuali di incremento e riduzione per l'utilizzo della CPU, determinate dal parametro `CpuUtilization`. Quando il parametro `CpuUtilization` del connettore supera la percentuale di dimensionamento orizzontale, MSK Connect aumenta il numero di worker che utilizzano il connettore. Quando il parametro `CpuUtilization` scende al di sotto della percentuale di ridimensionamento, MSK Connect riduce il numero di worker. Il numero di worker rimane sempre compreso tra il numero minimo e massimo specificato al momento della creazione del connettore.
  - Il numero di MCU per worker.

Per ulteriori informazioni sui worker, consulta la pagina the section called “Worker”. Per ulteriori informazioni sui parametri di MSK Connect, consulta la pagina the section called “Monitoraggio”.

Creazione di un connettore

Creazione di un connettore utilizzando AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Nel riquadro a sinistra, in MSK Connect, scegli Connettori.
3. Scegli Create connector (Crea connettore).
4. Per creare il connettore, puoi scegliere se utilizzare un plug-in personalizzato esistente o creare innanzitutto un nuovo plug-in personalizzato. Per informazioni sui plug-in personalizzati e su come crearli, consulta la pagina the section called “Plug-in”. In questa procedura,

supponiamo che tu abbia un plug-in personalizzato che desideri utilizzare. Nell'elenco dei plug-in personalizzati, trova quello che desideri utilizzare, seleziona la casella alla sua sinistra, quindi scegli Avanti.

5. Inserisci un nome e facoltativamente una descrizione.
6. Scegli il cluster a cui desideri connetterti.
7. Specifica la configurazione del connettore. I parametri di configurazione da specificare dipendono dal tipo di connettore che si desidera creare. Tuttavia, alcuni parametri sono comuni a tutti i connettori, ad esempio i parametri `connector.class` e `tasks.max`. Di seguito è riportato un esempio di configurazione per il connettore sink Amazon S3 Confluent.

```
connector.class=io.confluent.connect.s3.S3SinkConnector
tasks.max=2
topics=my-example-topic
s3.region=us-east-1
s3.bucket.name=my-destination-bucket
flush.size=1
storage.class=io.confluent.connect.s3.storage.S3Storage
format.class=io.confluent.connect.s3.format.json.JsonFormat
partitioner.class=io.confluent.connect.storage.partitioner.DefaultPartitioner
key.converter=org.apache.kafka.connect.storage.StringConverter
value.converter=org.apache.kafka.connect.storage.StringConverter
schema.compatibility=NONE
```

8. Successivamente, configura la capacità del connettore. È possibile scegliere tra due modalità di capacità: assegnata e con dimensionamento automatico. Per informazioni su queste due opzioni, consulta the section called “Capacità”.
9. Scegli la configurazione del worker predefinita o una configurazione del worker personalizzata. Per informazioni sulla creazione di configurazioni del worker personalizzate, consulta la pagina the section called “Worker”.
10. Successivamente, specifica il ruolo di esecuzione del servizio. Questo deve essere un ruolo IAM che MSK Connect può assumere e che concede al connettore tutte le autorizzazioni necessarie per accedere alle risorse necessarie. AWS Tali autorizzazioni dipendono dalla logica del connettore. Per informazioni su come creare questo ruolo, consulta the section called “Ruolo di esecuzione del servizio”.
11. Scegli Avanti, esamina le informazioni di sicurezza, quindi scegli nuovamente Avanti.
12. Specifica le opzioni di registrazione desiderate, quindi scegli Avanti. Per ulteriori informazioni sulla registrazione, consulta the section called “Registrazione”.

13. Scegli Create connector (Crea connettore).

Per utilizzare l'API MSK Connect per creare un connettore, vedere CreateConnector.

Plug-in

Un plugin è una AWS risorsa che contiene il codice che definisce la logica del connettore. Quando si crea il plug-in, si carica un file JAR (o un file ZIP che contiene uno o più file JAR) in un bucket S3 e si specifica la posizione del bucket. Quando si crea un connettore, si specifica il plug-in che si desidera che MSK Connect utilizzi. La relazione tra i plugin e i connettori è one-to-many: È possibile creare uno o più connettori dallo stesso plugin.

Per informazioni su come sviluppare il codice per un connettore, consulta la pagina Connector Development Guide nella documentazione di Apache Kafka.

Creazione di un plug-in personalizzato utilizzando il AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Nel riquadro a sinistra, in MSK Connect, scegli Plug-in personalizzati.
3. Scegli Crea plug-in personalizzato.
4. Seleziona Sfoglia S3.
5. Nell'elenco dei bucket S3, scegli il bucket contenente il file JAR o ZIP per il plug-in.
6. Nell'elenco degli oggetti, seleziona la casella a sinistra del file JAR o ZIP per il plug-in, quindi seleziona Scegli.
7. Scegli Crea plug-in personalizzato.

Per utilizzare l'API MSK Connect per creare un plug-in personalizzato, vedere CreateCustomPlugin.

Worker

Un worker è un processo di macchina virtuale Java (JVM) che esegue la logica del connettore. Ogni worker crea una serie di attività che vengono eseguite in thread paralleli e si occupano di copiare i dati. Le attività non memorizzano lo stato e possono quindi essere avviate, interrotte o riavviate in qualsiasi momento per fornire una pipeline di dati resiliente e scalabile. Le modifiche al numero di worker, dovute a un evento di dimensionamento o a guasti imprevisti, vengono rilevate automaticamente dai worker rimanenti. Essi si coordinano per riequilibrare le attività tra il gruppo

di worker rimanenti. I worker di Connect utilizzano i gruppi di consumatori di Apache Kafka per tali operazioni di coordinamento e riequilibrio.

Se i requisiti di capacità del connettore sono variabili o difficili da stimare, è possibile consentire a MSK Connect di dimensionare il numero di worker in base alle esigenze entro un limite inferiore e un limite superiore specificati. In alternativa, è possibile specificare il numero esatto di worker da utilizzare per l'esecuzione della logica di connessione. Per ulteriori informazioni, consulta the section called “Capacità”.

Gli operatori di MSK Connect utilizzano gli indirizzi IP

Gli operatori di MSK Connect utilizzano gli indirizzi IP nelle sottoreti fornite dal cliente. Ogni lavoratore utilizza un indirizzo IP da una delle sottoreti fornite dal cliente. È necessario assicurarsi di disporre di un numero sufficiente di indirizzi IP disponibili nelle sottoreti fornite a una CreateConnector richiesta per tenere conto della capacità specificata, specialmente quando si scalano automaticamente i connettori in cui il numero di lavoratori può variare.

Argomenti

Configurazione dei worker predefinita

MSK Connect fornisce la seguente configurazione predefinita per i worker:

```
key.converter=org.apache.kafka.connect.storage.StringConverter
value.converter=org.apache.kafka.connect.storage.StringConverter
```

Proprietà di configurazione dei worker supportate

MSK Connect fornisce una configurazione predefinita per i worker. Se lo desideri, puoi creare una configurazione dei worker personalizzata da utilizzare con i connettori. L'elenco seguente include informazioni sulle proprietà di configurazione dei worker supportate o meno da Amazon MSK Connect.

- Sono obbligatorie solo le proprietà `key.converter` e `value.converter`.

- MSK Connect supporta le seguenti proprietà di configurazione di `producer.`.

```
producer.acks
producer.batch.size
producer.buffer.memory
producer.compression.type
producer.enable.idempotence
producer.key.serializer
producer.max.request.size
producer.metadata.max.age.ms
producer.metadata.max.idle.ms
producer.partitioner.class
producer.reconnect.backoff.max.ms
producer.reconnect.backoff.ms
producer.request.timeout.ms
producer.retry.backoff.ms
producer.value.serializer
```

- MSK Connect supporta le seguenti proprietà di configurazione di `consumer.`.

```
consumer.allow.auto.create.topics
consumer.auto.offset.reset
consumer.check.crcs
consumer.fetch.max.bytes
consumer.fetch.max.wait.ms
consumer.fetch.min.bytes
consumer.heartbeat.interval.ms
consumer.key.deserializer
consumer.max.partition.fetch.bytes
consumer.max.poll.records
consumer.metadata.max.age.ms
consumer.partition.assignment.strategy
consumer.reconnect.backoff.max.ms
consumer.reconnect.backoff.ms
consumer.request.timeout.ms
consumer.retry.backoff.ms
consumer.session.timeout.ms
consumer.value.deserializer
```

- Sono supportate tutte le altre proprietà di configurazione che non iniziano con i prefissi `producer.` o `consumer.`, ad eccezione delle seguenti proprietà.

```
access.control.
```

```
admin.
admin.listeners.https.
client.
connect.
inter.worker.
internal.
listeners.https.
metrics.
metrics.context.
rest.
sasl.
security.
socket.
ssl.
topic.tracking.
worker.
bootstrap.servers
config.storage.topic
connections.max.idle.ms
connector.client.config.override.policy
group.id
listeners
metric.reporters
plugin.path
receive.buffer.bytes
response.http.headers.config
scheduled.rebalance.max.delay.ms
send.buffer.bytes
status.storage.topic
```

Per ulteriori informazioni sulle proprietà di configurazione dei worker e su cosa rappresentano, consulta la pagina Kafka Connect Configs nella documentazione di Apache Kafka.

Creazione di una configurazione dei worker personalizzata

Creazione di una configurazione di lavoro personalizzata utilizzando AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Nel riquadro a sinistra, in MSK Connect, scegli Configurazioni del worker.
3. Seleziona Configurazione del worker.

4. Inserisci un nome e una descrizione opzionale, quindi aggiungi le proprietà e i valori su cui desideri impostarli.
5. Seleziona Configurazione del worker.

Per utilizzare l'API MSK Connect per creare una configurazione del lavoratore, vedere CreateWorkerConfiguration.

Gestione degli offset dei connettori di origine tramite `offset.storage.topic`

Questa sezione fornisce informazioni per aiutarti a gestire gli offset dei connettori di origine tramite l'argomento di archiviazione degli offset. L'argomento di archiviazione degli offset è un argomento interno che Kafka Connect utilizza per archiviare gli offset di configurazione dei connettori e delle attività.

Utilizzo dell'argomento predefinito per l'archiviazione degli offset

Per impostazione predefinita, Amazon MSK Connect genera un nuovo argomento di archiviazione degli offset sul cluster Kafka per ogni connettore creato. MSK costruisce il nome dell'argomento predefinito utilizzando parti dell'ARN del connettore. Ad esempio, `__amazon_msk_connect_offsets_my-mskc-connector_12345678-09e7-4abc-8be8-c657f7e4ff32-2`.

Definizione di un argomento personalizzato per l'archiviazione degli offset

Per garantire la continuità degli offset tra i connettori di origine, puoi utilizzare un argomento di archiviazione degli offset a tua scelta anziché l'argomento predefinito. La definizione di un argomento di archiviazione degli offset consente di eseguire attività come la creazione di un connettore di origine che riprenda la lettura dall'ultimo offset di un connettore precedente.

Per definire un argomento di archiviazione degli offset, è necessario fornire un valore per la proprietà `offset.storage.topic` nella configurazione del worker prima di creare un connettore. Se si desidera riutilizzare l'argomento di archiviazione degli offset per utilizzare gli offset di un connettore creato in precedenza, è necessario assegnare al nuovo connettore lo stesso nome di quello precedente. Se si crea un argomento di archiviazione degli offset personalizzato, è necessario impostare `cleanup.policy` su `compact` nella configurazione dell'argomento.

> **Note**
>
> Se si specifica un argomento di archiviazione degli offset quando si crea un connettore sink, MSK Connect crea l'argomento, se non esiste ancora. Tuttavia, l'argomento non verrà utilizzato per archiviare gli offset dei connettori.
> Gli offset dei connettori sink vengono invece gestiti utilizzando il protocollo del gruppo di consumatori di Kafka. Ogni connettore sink crea un gruppo denominato `connect-{CONNECTOR_NAME}`. Finché esiste il gruppo di consumatori, tutti i connettori sink creati successivamente con lo stesso valore di `CONNECTOR_NAME` continueranno dall'ultimo offset confermato.

Example : definizione di un argomento di archiviazione degli offset per ricreare un connettore di origine con una configurazione aggiornata

Supponi di avere un connettore Change Data Capture (CDC) e di voler modificare la configurazione del connettore senza perdere il posto nel flusso CDC. Non è possibile aggiornare la configurazione esistente del connettore, ma è possibile eliminare il connettore e crearne uno nuovo con lo stesso nome. Per indicare al nuovo connettore da dove iniziare a leggere il flusso CDC, puoi specificare l'argomento di archiviazione degli offset del vecchio connettore nella configurazione del worker. Di seguito viene illustrato come realizzare tale operazione.

1. Sul computer client, esegui il comando seguente per trovare il nome dell'argomento archiviazione degli offset del connettore. Sostituisci *<bootstrapBrokerString>* con la stringa del broker di bootstrap del cluster. Per istruzioni su come recuperare la stringa del broker di bootstrap, consulta la pagina Recupero dei broker di bootstrap per un cluster Amazon MSK.

   ```
   <path-to-your-kafka-installation>/bin/kafka-topics.sh --list --bootstrap-
   server <bootstrapBrokerString>
   ```

   L'output seguente mostra un elenco di tutti gli argomenti del cluster, inclusi gli argomenti predefiniti relativi ai connettori interni. In questo esempio, il connettore CDC esistente utilizza l'argomento di archiviazione degli offset predefinito creato da MSK Connect. Questo è il motivo per cui l'argomento di archiviazione degli offset è chiamato `__amazon_msk_connect_offsets_my-mskc-connector_12345678-09e7-4abc-8be8-c657f7e4ff32-2`.

   ```
   __consumer_offsets
   ```

```
__amazon_msk_canary
__amazon_msk_connect_configs_my-mskc-connector_12345678-09e7-4abc-8be8-
c657f7e4ff32-2
__amazon_msk_connect_offsets_my-mskc-connector_12345678-09e7-4abc-8be8-
c657f7e4ff32-2
__amazon_msk_connect_status_my-mskc-connector_12345678-09e7-4abc-8be8-
c657f7e4ff32-2
my-msk-topic-1
my-msk-topic-2
```

2. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
3. Scegli il connettore dall'elenco Connettori. Copia e salva il contenuto del campo Configurazione del connettore in modo da poterlo modificare e utilizzare per creare il nuovo connettore.
4. Scegli Elimina per confermare l'eliminazione. Inserisci il nome del connettore nel campo di immissione del testo per confermare l'eliminazione.
5. Crea una configurazione di worker personalizzata con valori adatti al tuo scenario. Per istruzioni, consulta Creazione di una configurazione dei worker personalizzata.

   Nella configurazione del worker, è necessario specificare il nome dell'argomento di archiviazione degli offset recuperato in precedenza come valore di `offset.storage.topic`, come nella configurazione seguente.

   ```
   config.providers.secretManager.param.aws.region=us-east-1
   key.converter=<org.apache.kafka.connect.storage.StringConverter>
   value.converter=<org.apache.kafka.connect.storage.StringConverter>
   config.providers.secretManager.class=com.github.jcustenborder.kafka.config.aws.SecretsManag
   config.providers=secretManager
   offset.storage.topic=__amazon_msk_connect_offsets_my-mskc-
   connector_12345678-09e7-4abc-8be8-c657f7e4ff32-2
   ```

6. **Important**

   Al nuovo connettore deve essere assegnato lo stesso nome del vecchio connettore.

   Crea un nuovo connettore utilizzando la configurazione del worker impostata nel passaggio precedente. Per istruzioni, consulta Creazione di un connettore.

Considerazioni

Durante la gestione degli offset del connettore di origine, tieni in considerazione i seguenti aspetti.

- Per specificare un argomento archiviazione degli offset, fornisci il nome dell'argomento Kafka in cui gli offset dei connettori vengono archiviati come valore di `offset.storage.topic` nella configurazione del worker.
- Presta attenzione quando apporti modifiche alla configurazione di un connettore. Se un connettore di origine utilizza i valori della configurazione per record di offset chiave, la modifica dei valori di configurazione può causare un comportamento indesiderato del connettore. Ti consigliamo di fare riferimento alla documentazione del tuo plug-in per informazioni.
- Personalizza il numero predefinito di partizioni: oltre a personalizzare la configurazione del worker mediante l'aggiunta di `offset.storage.topic`, è possibile personalizzare il numero di partizioni per gli argomenti di archiviazione degli offset e degli stati. Le partizioni predefinite per gli argomenti interni sono le seguenti.
  - `config.storage.topic`: 1, non configurabile, deve essere un argomento a partizione singola
  - `offset.storage.topic`: 25, configurabile fornendo `offset.storage.partitions`
  - `status.storage.topic`: 5, configurabile fornendo `status.storage.partitions`
- Eliminazione manuale degli argomenti: Amazon MSK Connect crea nuovi argomenti interni di Kafka Connect (il nome dell'argomento inizia con `__amazon_msk_connect`) a ogni implementazione di connettori. I vecchi argomenti associati ai connettori eliminati non vengono rimossi automaticamente perché gli argomenti interni, ad esempio `offset.storage.topic`, possono essere riutilizzati tra i connettori. Tuttavia, è possibile eliminare manualmente gli argomenti interni non utilizzati creati da MSK Connect. Gli argomenti interni sono denominati secondo il formato `__amazon_msk_connect_<offsets|status|configs>_`*`connector_name_connector_id`*.

  Per eliminare gli argomenti interni, è possibile utilizzare l'espressione regolare `__amazon_msk_connect_<offsets|status|configs>_`*`connector_name_connector_id`*. Evita di eliminare un argomento interno attualmente utilizzato da un connettore in esecuzione.
- Utilizzo dello stesso nome per gli argomenti interni creati d MSK Connect: se desideri riutilizzare l'argomento di archiviazione degli offset per utilizzare gli offset di un connettore creato in precedenza, dovrai assegnare al nuovo connettore lo stesso nome di quello precedente. Nella configurazione del worker, è possibile impostare la proprietà `offset.storage.topic` per assegnare lo stesso nome a `offset.storage.topic` e riutilizzarlo tra connettori

diversi. Questa configurazione è descritta nella sezione Gestione degli offset dei connettori. MSK Connect non consente a connettori diversi di condividere `config.storage.topic` e `status.storage.topic`. Questi argomenti vengono creati ogni volta che si crea un nuovo connettore in MSKC. Vengono denominati automaticamente secondo il formato `__amazon_msk_connect_<status|configs>_`*`connector_name_connector_id`* e quindi sono diversi per ciascuno dei connettori creati.

Esternalizzazione di informazioni sensibili utilizzando i provider di configurazione

Questo esempio mostra come esternalizzare le informazioni sensibili per Amazon MSK Connect utilizzando un provider di configurazione open source. Un provider di configurazione consente di specificare variabili anziché testo non crittografato in una configurazione di connettore o di worker e i worker in esecuzione nel connettore risolvono queste variabili in fase di runtime. Ciò impedisce che le credenziali e altri segreti vengano archiviati in testo non crittografato. Il provider di configurazione nell'esempio supporta il recupero dei parametri di configurazione da AWS Secrets Manager, Amazon S3 e Systems Manager (SSM). Nel passaggio 2, viene illustrato come configurare l'archiviazione e il recupero di informazioni sensibili per il servizio che desideri configurare.

Argomenti

- Passaggio 1: creazione di un plug-in personalizzato e caricamento dello stesso su S3
- Passaggio 2: configurazione dei parametri e delle autorizzazioni per diversi provider
- Passaggio 3: creazione di una configurazione del worker personalizzata con informazioni sul proprio provider di configurazione
- Passaggio 4: creazione del connettore
- Considerazioni

Passaggio 1: creazione di un plug-in personalizzato e caricamento dello stesso su S3

Per creare un plugin personalizzato, crea un file zip che contenga il connettore ed esegui i seguenti comandi sul msk-config-provider tuo computer locale.

Creazione di un plug-in personalizzato utilizzando una finestra di terminale e Debezium come connettore

Usa la AWS CLI per eseguire comandi come superutente con credenziali che ti consentono di accedere al tuo bucket S3. AWS Per informazioni sull'installazione e la configurazione della AWS CLI, consulta Guida introduttiva alla AWS CLI nella Guida per l'utente.AWS Command Line Interface Per informazioni sull'uso della AWS CLI con Amazon S3, consulta Using Amazon S3 with the AWS CLI nella Guida per l'utente.AWS Command Line Interface

1. In una finestra del terminale, crea una cartella denominata `custom-plugin` nel tuo spazio di lavoro tramite il seguente comando.

   ```
   mkdir custom-plugin && cd custom-plugin
   ```

2. Scarica l'ultima versione stabile del plug-in per il connettore MySQL dal sito di Debezium tramite il seguente comando.

   ```
   wget https://repo1.maven.org/maven2/io/debezium/debezium-connectormysql/
   2.2.0.Final/debezium-connector-mysql-2.2.0.Final-plugin.tar.gz
   ```

   Estrai il file gzip scaricato nella cartella `custom-plugin` tramite il seguente comando.

   ```
   tar xzf debezium-connector-mysql-2.2.0.Final-plugin.tar.gz
   ```

3. Scarica il file zip del provider di configurazione MSK tramite il seguente comando.

   ```
   wget https://github.com/aws-samples/msk-config-providers/releases/download/r0.1.0/
   msk-config-providers-0.1.0-with-dependencies.zip
   ```

   Estrai il file zip scaricato nella cartella `custom-plugin` tramite il seguente comando.

   ```
   unzip msk-config-providers-0.1.0-with-dependencies.zip
   ```

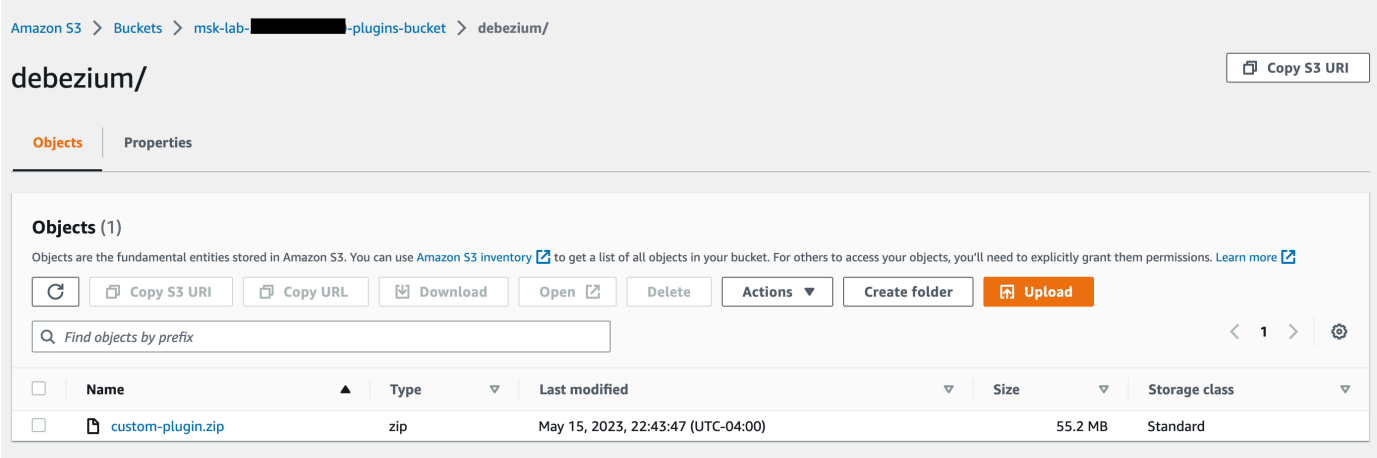
4. Comprimi il contenuto del provider di configurazione MSK del passaggio precedente e del connettore personalizzato in un unico file denominato `custom-plugin.zip`.

   ```
   zip -r ../custom-plugin.zip *
   ```

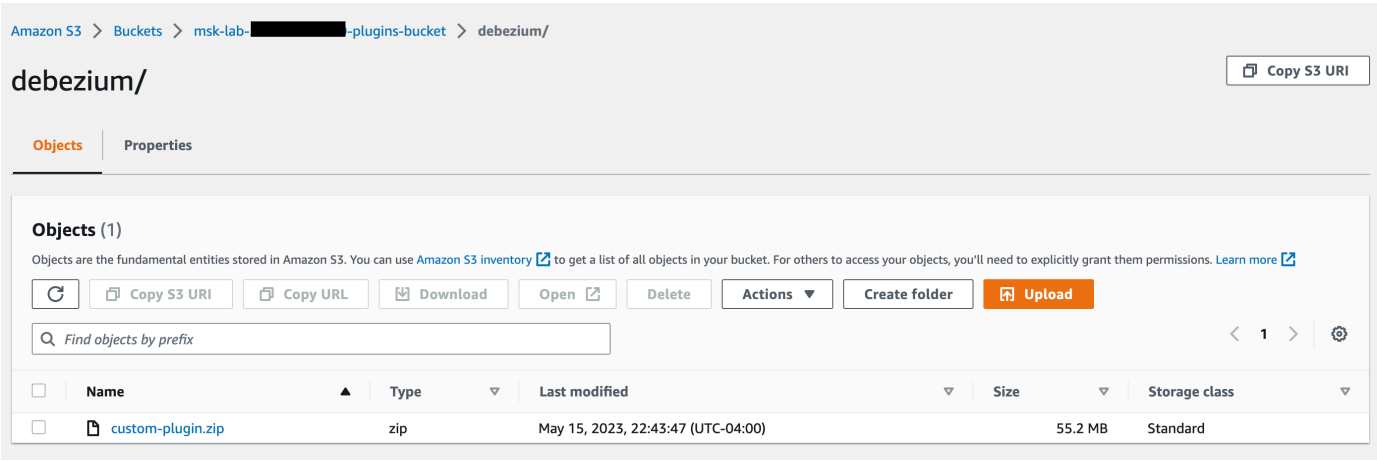
5. Carica il file su S3 per utilizzarlo come riferimento in seguito.

```
aws s3 cp ../custom-plugin.zip s3:<S3_URI_BUCKET_LOCATION>
```

6. Sulla console Amazon MSK, nella sezione MSK Connect, scegli Plug-in personalizzato, quindi scegli Crea plug-in personalizzato e sfoglia il bucket S3 s3:<***S3_URI_BUCKET_LOCATION***> per selezionare il file ZIP del plug-in personalizzato che hai appena caricato.



7. Inserisci **`debezium-custom-plugin`** come nome del plug-in. Facoltativamente, inserisci una descrizione e scegli Crea plug-in personalizzato.



Passaggio 2: configurazione dei parametri e delle autorizzazioni per diversi provider

È possibile configurare i valori dei parametri in questi tre servizi:

- Secrets Manager
- Systems Manager Parameter Store

- S3 - Simple Storage Service

Seleziona una delle schede seguenti per ottenere istruzioni sulla configurazione dei parametri e delle autorizzazioni pertinenti per tale servizio.

Configure in Secrets Manager

Configurazione dei valori dei parametri in Secrets Manager

1. Apri la console Secrets Manager.
2. Crea un nuovo segreto per archiviare le credenziali o i segreti. Per istruzioni, consulta Creare un AWS Secrets Manager segreto nella Guida per l'utente. AWS Secrets Manager
3. Copia l'ARN del segreto.
4. Aggiungi le autorizzazioni di Secrets Manager dalla seguente policy di esempio al tuo ruolo di esecuzione del servizio. Sostituisci *<arn:aws:secretsmanager:us-east-1:123456789000:secret*: -1234> con l'ARN del tuo segreto. MySecret
5. Aggiungi le istruzioni per la configurazione del worker e il connettore.

```
{
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "secretsmanager:GetResourcePolicy",
                    "secretsmanager:GetSecretValue",
                    "secretsmanager:DescribeSecret",
                    "secretsmanager:ListSecretVersionIds"
                ],
                "Resource": [
                "<arn:aws:secretsmanager:us-
east-1:123456789000:secret:MySecret-1234>"
                ]
            }
        ]
    }
```

6. Per utilizzare il provider di configurazione Secrets Manager, copia le seguenti righe di codice nella casella di testo della configurazione del worker nel passaggio 3:

```
# define name of config provider:

config.providers = secretsmanager

# provide implementation classes for secrets manager:

config.providers.secretsmanager.class =
 com.amazonaws.kafka.config.providers.SecretsManagerConfigProvider

# configure a config provider (if it needs additional initialization), for
 example you can provide a region where the secrets or parameters are located:

config.providers.secretsmanager.param.region = us-east-1
```

7. Per il provider di configurazione Secrets Manager, copia le seguenti righe di codice nella configurazione del connettore nel passaggio 4.

```
#Example implementation for secrets manager variable
database.hostname=${secretsmanager:MSKAuroraDBCredentials:username}

database.password=${secretsmanager:MSKAuroraDBCredentials:password}
```

È possibile utilizzare il passaggio precedente anche con altri provider di configurazione.

Configure in Systems Manager Parameter Store

Configurazione dei valori dei parametri in Archivio dei parametri Systems Manager

1. Aprire la console Systems Manager.
2. Nel riquadro di navigazione, selezionare Parameter Store (Archivio parametri).
3. Crea un nuovo parametro da archiviare in Systems Manager. Per istruzioni, vedere Create a Systems Manager (console) nella Guida per l' AWS Systems Manager utente.
4. Copia l'ARN del parametro.
5. Aggiungi le autorizzazioni di Systems Manager dalla seguente policy di esempio al tuo ruolo di esecuzione del servizio. Sostituisci *<arn:aws:ssm:us-east- MyParameterName 1:123456789000:parameter/* > con l'ARN del tuo parametro.

```
{
        "Version": "2012-10-17",
```

```
        "Statement": [
            {
                "Sid": "VisualEditor0",
                "Effect": "Allow",
                "Action": [
                    "ssm:GetParameterHistory",
                    "ssm:GetParametersByPath",
                    "ssm:GetParameters",
                    "ssm:GetParameter"
                ],
                "Resource": "arn:aws:ssm:us-east-1:123456789000:parameter/
MyParameterName"
            }
        ]
    }
```

6. Per utilizzare il provider di configurazione Archivio dei parametri, copia le seguenti righe di codice nella casella di testo della configurazione del worker nel passaggio 3:

```
# define name of config provider:

config.providers = ssm

# provide implementation classes for parameter store:

config.providers.ssm.class =
 com.amazonaws.kafka.config.providers.SsmParamStoreConfigProvider

# configure a config provider (if it needs additional initialization), for
 example you can provide a region where the secrets or parameters are located:

config.providers.ssm.param.region = us-east-1
```

7. Per il provider di configurazione Archivio dei parametri, copia le seguenti righe di codice nella configurazione del connettore nel passaggio 5.

```
#Example implementation for parameter store variable
schema.history.internal.kafka.bootstrap.servers=
${ssm::MSKBootstrapServerAddress}
```

È possibile utilizzare i due passaggi precedenti anche con altri provider di configurazione.

Configure in Amazon S3

Configurazione di oggetti/file in Amazon S3

1. Apri la console Amazon S3.
2. Carica l'oggetto in un bucket S3. Per istruzioni, consulta la pagina Uploading objects.
3. Copia l'ARN dell'oggetto.
4. Aggiungi le autorizzazioni di Amazon S3 Object Read dalla seguente policy di esempio al tuo ruolo di esecuzione del servizio. Sostituisci *<arn:aws:s3:::MY_S3_BUCKET/path/to/custom-plugin.zip>* con l'ARN dell'oggetto.

```
{
    "Version": "2012-10-17",
    "Statement": [
            {
                "Sid": "VisualEditor0",
                "Effect": "Allow",
                "Action": "s3:GetObject",
                "Resource": "<arn:aws:s3:::MY_S3_BUCKET/path/to/custom-
plugin.zip>"
            }
        ]
    }
```

5. Per utilizzare il provider di configurazione Amazon S3, copia le seguenti righe di codice nella casella di testo della configurazione del worker nel passaggio 3:

```
# define name of config provider:

config.providers = s3import
# provide implementation classes for S3:

config.providers.s3import.class =
 com.amazonaws.kafka.config.providers.S3ImportConfigProvider
```

6. Per il provider di configurazione Amazon S3, copia le seguenti righe di codice nella configurazione del connettore nel passaggio 4.

```
#Example implementation for S3 object
```

```
database.ssl.truststore.location = ${s3import:us-west-2:my_cert_bucket/path/to/
trustore_unique_filename.jks}
```

È possibile utilizzare i due passaggi precedenti anche con altri provider di configurazione.

Passaggio 3: creazione di una configurazione del worker personalizzata con informazioni sul proprio provider di configurazione

1. Seleziona Configurazioni dei worker nella sezione Amazon MSK Connect.
2. Seleziona Crea configurazione del worker.
3. Inserisci `SourceDebeziumCustomConfig` nella casella di testo Nome della configurazione del worker. La descrizione è facoltativa.
4. Copia il codice di configurazione pertinente in base ai provider desiderati e incollalo nella casella di testo Configurazione del worker.
5. Questo è un esempio di configurazione del worker per tutti e tre i provider:

```
key.converter=org.apache.kafka.connect.storage.StringConverter
key.converter.schemas.enable=false
value.converter=org.apache.kafka.connect.json.JsonConverter
value.converter.schemas.enable=false
offset.storage.topic=offsets_my_debezium_source_connector

# define names of config providers:

config.providers=secretsmanager,ssm,s3import

# provide implementation classes for each provider:

config.providers.secretsmanager.class    =
 com.amazonaws.kafka.config.providers.SecretsManagerConfigProvider
config.providers.ssm.class               =
 com.amazonaws.kafka.config.providers.SsmParamStoreConfigProvider
config.providers.s3import.class          =
 com.amazonaws.kafka.config.providers.S3ImportConfigProvider

# configure a config provider (if it needs additional initialization), for example
 you can provide a region where the secrets or parameters are located:
```

```
config.providers.secretsmanager.param.region = us-east-1
config.providers.ssm.param.region = us-east-1
```

6. Fai clic su Crea configurazione del worker.

Passaggio 4: creazione del connettore

1. Crea un nuovo connettore seguendo le istruzioni riportate nella sezione Creazione di un nuovo connettore.
2. Scegli il file `custom-plugin.zip` che hai caricato nel tuo bucket S3 in ??? come origine del plug-in personalizzato.
3. Copia il codice di configurazione pertinente in base ai provider desiderati e incollalo nel campo Configurazione del cluster.
4. Questo è un esempio della configurazione dei connettori per tutti e tre i provider:

```
#Example implementation for parameter store variable
schema.history.internal.kafka.bootstrap.servers=${ssm::MSKBootstrapServerAddress}

#Example implementation for secrets manager variable
database.hostname=${secretsmanager:MSKAuroraDBCredentials:username}

database.password=${secretsmanager:MSKAuroraDBCredentials:password}

#Example implementation for Amazon S3 file/object
database.ssl.truststore.location = ${s3import:us-west-2:my_cert_bucket/path/to/
trustore_unique_filename.jks}
```

5. Seleziona Usa una configurazione personalizzata e scegli dal menu a discesa Worker Configuration SourceDebeziumCustomConfig.
6. Segui i passaggi rimanenti indicati nelle istruzioni nella sezione Creazione di un connettore.

Considerazioni

Considera quanto segue durante l'utilizzo del provider di configurazione MSK con Amazon MSK Connect:

- Quando utilizzi i provider di configurazione, assegna le autorizzazioni appropriate al ruolo di esecuzione del servizio IAM.

- Definisci i provider di configurazione nelle configurazioni dei worker e la rispettiva implementazione nella configurazione del connettore.
- Se un plug-in non definisce i valori di configurazione sensibili come segreti, tali valori possono apparire nei log dei connettori. Kafka Connect tratta i valori di configurazione non definiti allo stesso modo di qualsiasi altro valore non crittografato. Per ulteriori informazioni, consulta Impedire la visualizzazione di segreti nei log dei connettori.
- Per impostazione predefinita, spesso MSK Connect riavvia un connettore se questo utilizza un provider di configurazione. Per disattivare questo comportamento di riavvio, è possibile impostare il valore di `config.action.reload` su `none` nella configurazione del connettore.

Ruoli e policy IAM per MSK Connect

Argomenti

- Ruolo di esecuzione del servizio
- Esempi di policy IAM per MSK Connect
- Prevenzione del confused deputy tra servizi
- AWS politiche gestite per MSK Connect
- Utilizzo dei ruoli collegati ai servizi per MSK Connect

Ruolo di esecuzione del servizio

> **Note**
>
> Amazon MSK Connect non supporta l'utilizzo del ruolo collegato ai servizi come ruolo di esecuzione del servizio. È necessario creare un ruolo di esecuzione del servizio separato. Per istruzioni su come creare un ruolo IAM personalizzato, consulta Creare un ruolo per delegare le autorizzazioni a un AWS servizio nella Guida per l'utente IAM.

Quando si crea un connettore con MSK Connect, è necessario specificare un ruolo AWS Identity and Access Management (IAM) da utilizzare con esso. Il ruolo di esecuzione del servizio deve disporre della seguente policy di attendibilità affinché MSK Connect lo possa assumere. Per ulteriori informazioni sulle chiavi di contesto delle condizioni in questa policy, consulta la pagina the section called “Prevenzione del confused deputy tra servizi”.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "kafkaconnect.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "Account-ID"
        },
        "ArnLike": {
          "aws:SourceArn": "MSK-Connector-ARN"
        }
      }
    }
  ]
}
```

Se il cluster Amazon MSK che desideri utilizzare con il connettore è un cluster che utilizza l'autenticazione IAM, devi aggiungere la seguente policy di autorizzazione al ruolo di esecuzione del servizio del connettore. Per informazioni su come trovare l'UUID del cluster e su come costruire gli ARN di argomento, consulta la pagina the section called “Risorse”.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:Connect",
                "kafka-cluster:DescribeCluster"
            ],
            "Resource": [
                "cluster-arn"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
```

```
                "kafka-cluster:ReadData",
                "kafka-cluster:DescribeTopic"
            ],
            "Resource": [
                "ARN of the topic that you want a sink connector to read from"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:WriteData",
                "kafka-cluster:DescribeTopic"
            ],
            "Resource": [
                "ARN of the topic that you want a source connector to write to"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:CreateTopic",
                "kafka-cluster:WriteData",
                "kafka-cluster:ReadData",
                "kafka-cluster:DescribeTopic"
            ],
            "Resource": [
                "arn:aws:kafka:region:account-id:topic/cluster-name/cluster-uuid/
__amazon_msk_connect_*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:AlterGroup",
                "kafka-cluster:DescribeGroup"
            ],
            "Resource": [
                "arn:aws:kafka:region:account-id:group/cluster-name/cluster-uuid/
__amazon_msk_connect_*",
                "arn:aws:kafka:region:account-id:group/cluster-name/cluster-uuid/
connect-*"
            ]
        }
    ]
```

```
}
```

A seconda del tipo di connettore, potrebbe anche essere necessario allegare al ruolo di esecuzione del servizio una politica di autorizzazioni che gli consenta di accedere alle risorse. AWS Ad esempio, se il connettore deve inviare dati a un bucket S3, il ruolo di esecuzione del servizio deve disporre di una policy di autorizzazione che conceda l'autorizzazione alla scrittura su quel bucket. A scopo di test, puoi utilizzare una delle policy IAM predefinite che forniscono l'accesso completo, come `arn:aws:iam::aws:policy/AmazonS3FullAccess`. Tuttavia, per motivi di sicurezza, si consiglia di utilizzare la politica più restrittiva che consenta al connettore di leggere dalla AWS fonte o scrivere nel AWS sink.

Esempi di policy IAM per MSK Connect

Per fornire a un utente non amministratore l'accesso completo a tutte le funzionalità di MSK Connect, collega una policy come la seguente al ruolo IAM dell'utente.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kafkaconnect:*",
                "ec2:CreateNetworkInterface",
                "ec2:DescribeSubnets",
                "ec2:DescribeVpcs",
                "ec2:DescribeSecurityGroups",
                "logs:CreateLogDelivery",
                "logs:GetLogDelivery",
                "logs:DeleteLogDelivery",
                "logs:ListLogDeliveries",
                "logs:PutResourcePolicy",
                "logs:DescribeResourcePolicies",
                "logs:DescribeLogGroups"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": "iam:CreateServiceLinkedRole",
            "Resource": "arn:aws:iam::*:role/aws-service-role/
kafkaconnect.amazonaws.com/AWSServiceRoleForKafkaConnect*",
```

```
            "Condition": {
                "StringLike": {
                    "iam:AWSServiceName": "kafkaconnect.amazonaws.com"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "iam:AttachRolePolicy",
                "iam:PutRolePolicy"
            ],
            "Resource": "arn:aws:iam::*:role/aws-service-role/
kafkaconnect.amazonaws.com/AWSServiceRoleForKafkaConnect*"
        },
        {
            "Effect": "Allow",
            "Action": "iam:CreateServiceLinkedRole",
            "Resource": "arn:aws:iam::*:role/aws-service-role/
delivery.logs.amazonaws.com/AWSServiceRoleForLogDelivery*",
            "Condition": {
                "StringLike": {
                    "iam:AWSServiceName": "delivery.logs.amazonaws.com"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:PutBucketPolicy",
                "s3:GetBucketPolicy"
            ],
            "Resource": "ARN of the Amazon S3 bucket to which you want MSK Connect to
 deliver logs"
        },
        {
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": "ARN of the service execution role"
        },
        {
            "Effect": "Allow",
            "Action": "s3:GetObject",
```

```
            "Resource": "ARN of the Amazon S3 object that corresponds to the custom
 plugin that you want to use for creating connectors"
        },
        {
            "Effect": "Allow",
            "Action": "firehose:TagDeliveryStream",
            "Resource": "ARN of the Firehose delivery stream to which you want MSK
 Connect to deliver logs"
        }
    ]
}
```

Prevenzione del confused deputy tra servizi

Con "confused deputy" si intende un problema di sicurezza in cui un'entità che non dispone dell'autorizzazione per eseguire una certa operazione può costringere un'entità con più privilegi a eseguire tale operazione. Nel AWS, l'impersonificazione tra servizi può portare al confuso problema del vice. La rappresentazione tra servizi può verificarsi quando un servizio (il servizio chiamante) effettua una chiamata a un altro servizio (il servizio chiamato). Il servizio chiamante può essere manipolato per utilizzare le proprie autorizzazioni e agire sulle risorse di un altro cliente, a cui normalmente non avrebbe accesso. Per evitare ciò, AWS fornisce strumenti per poterti a proteggere i tuoi dati per tutti i servizi con entità di servizio a cui è stato concesso l'accesso alle risorse del tuo account.

Ti consigliamo di utilizzare le chiavi di contesto delle condizioni globali `aws:SourceArn` e `aws:SourceAccount` nelle policy delle risorse per limitare le autorizzazioni con cui MSK Connect fornisce un altro servizio alla risorsa. Se il valore `aws:SourceArn` non contiene l'ID account (ad esempio, l'AR di un bucket Amazon S3 non contiene l'ID account), è necessario utilizzare entrambe le chiavi di contesto delle condizioni globali per limitare le autorizzazioni. Se si utilizzano entrambe le chiavi di contesto delle condizioni globali e il valore `aws:SourceArn` contiene l'ID account, il valore `aws:SourceAccount` e l'account nel valore `aws:SourceArn` deve utilizzare lo stesso ID account nella stessa dichiarazione di policy. Utilizzare `aws:SourceArn` se si desidera consentire l'associazione di una sola risorsa all'accesso tra servizi. Utilizza `aws:SourceAccount` se desideri consentire l'associazione di qualsiasi risorsa in tale account all'uso tra servizi.

Nel caso di MSK Connect, il valore di `aws:SourceArn` deve essere un connettore MSK.

Il modo più efficace per proteggersi dal problema "confused deputy" è quello di usare la chiave di contesto della condizione globale `aws:SourceArn` con l'ARN completo della risorsa. Se non conosci

l'ARN completo della risorsa o scegli più risorse, utilizza la chiave di contesto della condizione globale `aws:SourceArn` con caratteri jolly (*) per le parti sconosciute dell'ARN. Ad esempio, *`arn:aws:kafkaconnect:us-east-1:123456789012:connector/*`* rappresenta tutti i connettori che appartengono all'account con ID 123456789012 nella regione Stati Uniti orientali (Virginia settentrionale).

L'esempio seguente mostra il modo in cui puoi utilizzare le chiavi di contesto delle condizioni globali `aws:SourceArn` e `aws:SourceAccount` in MSK Connect per prevenire il problema confused deputy. Sostituisci *`Account-ID`* e *`MSK-Connector-ARN`* con le tue informazioni.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": " kafkaconnect.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "Account-ID"
        },
        "ArnLike": {
          "aws:SourceArn": "MSK-Connector-ARN"
        }
      }
    }
  ]
}
```

AWS politiche gestite per MSK Connect

Una politica AWS gestita è una politica autonoma creata e amministrata da. AWS AWS le politiche gestite sono progettate per fornire autorizzazioni per molti casi d'uso comuni, in modo da poter iniziare ad assegnare autorizzazioni a utenti, gruppi e ruoli.

Tieni presente che le policy AWS gestite potrebbero non concedere le autorizzazioni con il privilegio minimo per i tuoi casi d'uso specifici, poiché sono disponibili per tutti i clienti. AWS Consigliamo pertanto di ridurre ulteriormente le autorizzazioni definendo policy gestite dal cliente specifiche per i tuoi casi d'uso.

Non è possibile modificare le autorizzazioni definite nelle politiche gestite. AWS Se AWS aggiorna le autorizzazioni definite in una politica AWS gestita, l'aggiornamento ha effetto su tutte le identità principali (utenti, gruppi e ruoli) a cui è associata la politica. AWS è più probabile che aggiorni una policy AWS gestita quando ne Servizio AWS viene lanciata una nuova o quando diventano disponibili nuove operazioni API per i servizi esistenti.

Per ulteriori informazioni, consultare Policy gestite da AWSnella Guida per l'utente di IAM.

AWS politica gestita: AmazonMSK ConnectReadOnlyAccess

Questa policy concede all'utente le autorizzazioni necessarie per elencare e descrivere le risorse MSK Connect.

È possibile allegare la policy `AmazonMSKConnectReadOnlyAccess`alle identità IAM.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kafkaconnect:ListConnectors",
                "kafkaconnect:ListCustomPlugins",
                "kafkaconnect:ListWorkerConfigurations"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafkaconnect:DescribeConnector"
            ],
            "Resource": [
                "arn:aws:kafkaconnect:*:*:connector/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafkaconnect:DescribeCustomPlugin"
            ],
            "Resource": [
                "arn:aws:kafkaconnect:*:*:custom-plugin/*"
```

```
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafkaconnect:DescribeWorkerConfiguration"
            ],
            "Resource": [
                "arn:aws:kafkaconnect:*:*:worker-configuration/*"
            ]
        }
    ]
}
```

AWS politica gestita: KafkaConnectServiceRolePolicy

Questa policy concede al servizio MSK Connect le autorizzazioni necessarie per creare e gestire le interfacce di rete alle quali è assegnato il tag `AmazonMSKConnectManaged:true`. Queste interfacce di rete forniscono a MSK Connect l'accesso di rete alle risorse del tuo Amazon VPC, come un cluster Apache Kafka, un'origine o un sink.

Non puoi collegarti KafkaConnectServiceRolePolicy alle tue entità IAM. Questa policy è collegata a un ruolo collegato ai servizi che consente a MSK Connect di eseguire operazioni per tuo conto.

```
{
 "Version": "2012-10-17",
 "Statement": [
  {
   "Effect": "Allow",
   "Action": [
    "ec2:CreateNetworkInterface"
   ],
   "Resource": "arn:aws:ec2:*:*:network-interface/*",
   "Condition": {
    "StringEquals": {
     "aws:RequestTag/AmazonMSKConnectManaged": "true"
    },
    "ForAllValues:StringEquals": {
     "aws:TagKeys": "AmazonMSKConnectManaged"
    }
   }
  },
  {
```

```
    "Effect": "Allow",
    "Action": [
     "ec2:CreateNetworkInterface"
    ],
    "Resource": [
     "arn:aws:ec2:*:*:subnet/*",
     "arn:aws:ec2:*:*:security-group/*"
    ]
   },
   {
    "Effect": "Allow",
    "Action": [
     "ec2:CreateTags"
    ],
    "Resource": "arn:aws:ec2:*:*:network-interface/*",
    "Condition": {
     "StringEquals": {
      "ec2:CreateAction": "CreateNetworkInterface"
     }
    }
   },
   {
    "Effect": "Allow",
    "Action": [
     "ec2:DescribeNetworkInterfaces",
     "ec2:CreateNetworkInterfacePermission",
     "ec2:AttachNetworkInterface",
     "ec2:DetachNetworkInterface",
     "ec2:DeleteNetworkInterface"
    ],
    "Resource": "arn:aws:ec2:*:*:network-interface/*",
    "Condition": {
     "StringEquals": {
      "ec2:ResourceTag/AmazonMSKConnectManaged": "true"
     }
    }
   }
  ]
}
```

MSK Connect aggiorna le policy AWS gestite

Visualizza i dettagli sugli aggiornamenti delle politiche AWS gestite per MSK Connect da quando questo servizio ha iniziato a tenere traccia di queste modifiche.

Modifica	Descrizione	Data
Policy di sola lettura di MSK Connect aggiornata	MSK Connect ha aggiornato la ConnectReadOnlyAccess politica di AmazonMSK per rimuovere le restrizioni sulle operazioni di pubblicazione delle offerte.	13 ottobre 2021
Inizio del tracciamento delle modifiche da parte di MSK Connect	MSK Connect ha iniziato a tenere traccia delle modifiche per le sue politiche AWS gestite.	14 settembre 2021

Utilizzo dei ruoli collegati ai servizi per MSK Connect

Amazon MSK Connect utilizza ruoli collegati ai servizi AWS Identity and Access Management (IAM). Un ruolo collegato ai servizi è un tipo di ruolo IAM univoco collegato direttamente a MSK Connect. I ruoli collegati ai servizi sono predefiniti da MSK Connect e includono tutte le autorizzazioni richieste dal servizio per chiamare altri AWS servizi per conto dell'utente.

Un ruolo collegato ai servizi semplifica la configurazione di MSK Connect perché permette di evitare l'aggiunta manuale delle autorizzazioni necessarie. MSK Connect definisce le autorizzazioni dei relativi ruoli associati ai servizi e, salvo diversamente definito, solo MSK Connect potrà assumere i propri ruoli. Le autorizzazioni definite includono la policy di attendibilità e la policy delle autorizzazioni che non può essere collegata a nessun'altra entità IAM.

Per informazioni sugli altri servizi che supportano i ruoli collegati ai servizi, consulta Servizi AWS che funzionano con IAM e cerca i servizi che riportano Sì nella colonna Ruolo associato ai servizi. Scegli Sì in corrispondenza di un link per visualizzare la documentazione relativa al ruolo collegato ai servizi per tale servizio.

Autorizzazioni del ruolo collegato ai servizi per MSK Connect

MSK Connect utilizza il ruolo collegato al servizio denominato: AWSServiceRoleForKafkaConnectconsente ad Amazon MSK Connect di accedere alle risorse Amazon per tuo conto.

Il ruolo AWSServiceRoleForKafkaConnect collegato al servizio si fida che il servizio assuma il ruolo. `kafkaconnect.amazonaws.com`

Per informazioni sulla policy di autorizzazione utilizzata dal ruolo, consulta la pagina the section called “KafkaConnectServiceRolePolicy”.

Per consentire a un'entità IAM (come un utente, un gruppo o un ruolo) di creare, modificare o eliminare un ruolo collegato ai servizi devi configurare le relative autorizzazioni. Per ulteriori informazioni, consulta Autorizzazioni del ruolo collegato ai servizi nella Guida per l'utente di IAM.

Creazione di un ruolo collegato ai servizi per MSK Connect

Non hai bisogno di creare manualmente un ruolo collegato ai servizi. Quando si crea un connettore nella AWS Management Console, la o l' AWS API AWS CLI, MSK Connect crea automaticamente il ruolo collegato al servizio.

Se elimini questo ruolo collegato ai servizi, puoi ricrearlo seguendo lo stesso processo utilizzato per ricreare il ruolo nell'account. Quando crei un connettore, MSK Connect crea di nuovo automaticamente il ruolo collegato ai servizi per conto dell'utente.

Modifica di un ruolo collegato ai servizi per MSK Connect

MSK Connect non consente di modificare il ruolo collegato al AWSServiceRoleForKafkaConnect servizio. Dopo aver creato un ruolo collegato al servizio, non puoi modificarne il nome, perché potrebbero farvi riferimento diverse entità. Puoi tuttavia modificarne la descrizione utilizzando IAM. Per ulteriori informazioni, consulta la sezione Modifica di un ruolo collegato ai servizi nella Guida per l'utente di IAM.

Eliminazione di un ruolo collegato ai servizi per MSK Connect

È possibile utilizzare la console IAM, AWS CLI o l' AWS API per eliminare manualmente il ruolo collegato al servizio. Per farlo, sarà necessario eliminare innanzitutto manualmente i connettori MSK Connect e quindi eliminare il ruolo manualmente. Per ulteriori informazioni, consulta Eliminazione del ruolo collegato ai servizi nella Guida per l'utente di IAM.

Regioni supportate per i ruoli collegati ai servizi di MSK Connect

MSK Connect supporta l'utilizzo di ruoli collegati ai servizi in tutte le regioni in cui il servizio è disponibile. Per ulteriori informazioni, consulta Regioni ed endpoint di AWS.

Abilitazione dell'accesso a Internet per Amazon MSK Connect

Se il tuo connettore per Amazon MSK Connect richiede l'accesso a Internet, ti consigliamo di utilizzare le seguenti impostazioni Amazon Virtual Private Cloud (VPC) per abilitare tale accesso.

- Configura il connettore con sottoreti private.
- Crea un gateway NAT pubblico o un'istanza NAT per il tuo VPC in una sottorete pubblica. Per ulteriori informazioni, consulta la pagina Connect subnet a Internet o ad altri VPC utilizzando dispositivi NAT nella Guida per l'utente. Amazon Virtual Private Cloud
- Consenti il traffico in uscita dalle sottoreti private verso il gateway o l'istanza NAT.

Configurazione di un gateway NAT per Amazon MSK Connect

Nei passaggi seguenti viene illustrato come configurare un gateway NAT per abilitare l'accesso a Internet per un connettore. È necessario completare questi passaggi prima di creare un connettore in una sottorete privata.

Prerequisiti

Verifica di disporre dei seguenti elementi.

- L'ID del Amazon Virtual Private Cloud (VPC) associato al cluster. Ad esempio, vpc-123456ab.
- Gli ID delle sottoreti private nel VPC. Ad esempio, subnet-a1b2c3de, subnet-f4g5h6ij e così via. Il connettore deve essere configurato con sottoreti private.

Abilitazione dell'accesso a Internet per il connettore

1. Apri la Amazon Virtual Private Cloud console all'indirizzo https://console.aws.amazon.com/vpc/.
2. Crea una sottorete pubblica con un nome descrittivo per il gateway NAT e prendi nota dell'ID della sottorete. Per istruzioni dettagliate, consulta la pagina Create a subnet in your VPC.

3. Crea un gateway Internet in modo che il VPC possa comunicare con Internet e prendi nota dell'ID del gateway. Collega il gateway Internet al VPC. Per istruzioni, consulta la pagina Create and attach an internet gateway.
4. Fornisci un gateway NAT pubblico in modo che gli host delle tue sottoreti private possano raggiungere la tua sottorete pubblica. Quando crei il gateway NAT, seleziona la sottorete pubblica creata in precedenza. Per istruzioni, consulta Creazione di un gateway NAT.
5. Configura le tabelle di routing. Per completare questa configurazione, occorrono in totale due tabelle di routing. Dovresti già disporre di una tabella di routing principale creata in automatico al momento della creazione del VPC. In questo passaggio creerai una tabella di routing aggiuntiva per la sottorete pubblica.

   a. Utilizza le seguenti impostazioni per modificare la tabella di routing principale del tuo VPC in modo che le sottoreti private instradino il traffico verso il tuo gateway NAT. Per le istruzioni, consulta la pagina Utilizzo delle tabelle di routing nella Guida per l'utente di Amazon Virtual Private Cloud.

      Tabella di routing MSKC privata

Proprietà	Valore
Name tag (Tag nome)	Ti consigliamo di assegnare a questa tabella di routing un nome descrittivo per facilitarne l'identificazione. Ad esempio, MSKC privata.
Sottoreti associate	Le tue sottoreti private
Un percorso per consentire l'accesso a Internet per MSK Connect	• Destinazione: 0.0.0.0/0 • Obiettivo: l'ID del gateway NAT. Ad esempio, nat-12a345bc6789efg1h.
Un percorso per tutto il traffico locale	• Destinazione: 10.0.0.0/16. Questo valore può variare a seconda del blocco CIDR del tuo VPC. • Obiettivo: locale

   b. Segui le istruzioni riportate nella pagina Creazione di una tabella di routing personalizzata per creare una tabella di routing per la sottorete pubblica. Quando crei la tabella, inserisci

un nome descrittivo nel campo Tag nome per identificare a quale sottorete è associata la tabella. Ad esempio, MSKC pubblica.

c. Configura la tua tabella di routing MSKC pubblica utilizzando le seguenti impostazioni.

Proprietà	Valore
Name tag (Tag nome)	MSKC pubblica o un altro nome descrittivo a scelta
Sottoreti associate	La tua sottorete pubblica con gateway NAT
Un percorso per consentire l'accesso a Internet per MSK Connect	• Destinazione: 0.0.0.0/0 • Obiettivo: l'ID del gateway Internet. Ad esempio, igw-1a234bc5.
Un percorso per tutto il traffico locale	• Destinazione: 10.0.0.0/16. Questo valore può variare a seconda del blocco CIDR del tuo VPC. • Obiettivo: locale

Nomi host DNS privati

Con il supporto dei nomi host Private DNS in MSK Connect, è possibile configurare i connettori per fare riferimento a nomi di dominio pubblici o privati. Il supporto dipende dai server DNS specificati nel set di opzioni DHCP del VPC.

Un set di opzioni DHCP è un gruppo di configurazioni di rete utilizzate dalle istanze EC2 nel VPC per comunicare tramite la rete VPC. Ogni VPC ha un set di opzioni DHCP predefinito ma è possibile creare un set di opzioni DHCP personalizzato se, ad esempio, si desidera che le istanze nel VPC utilizzino un server DNS diverso per la risoluzione dei nomi di dominio anziché il server DNS fornito da Amazon. Consulta la pagina DHCP option sets in Amazon VPC.

Prima che la funzionalità di risoluzione Private DNS fosse inclusa in MSK Connect, i connettori utilizzavano i risolutori DNS del servizio VPC per le query DNS provenienti da un connettore del cliente. I connettori non utilizzavano i server DNS definiti nei set di opzioni DHCP del VPC del cliente per la risoluzione DNS.

I connettori potevano fare riferimento solo ai nomi host nelle configurazioni dei connettori dei clienti o nei plug-in risolvibili pubblicamente. Non potevano risolvere nomi host privati definiti in una zona ospitata privatamente o utilizzare server DNS in una rete di altri clienti.

Senza Private DNS, i clienti che hanno scelto di rendere inaccessibili a Internet i propri database, data warehouse e sistemi come Secrets Manager nel proprio VPC, non potrebbero lavorare con i connettori MSK. I clienti utilizzano spesso nomi host DNS privati per conformarsi alle norme di sicurezza aziendali.

Argomenti

Configurazione di un set di opzioni DHCP del VPC per il connettore

I connettori utilizzano automaticamente i server DNS definiti nel set di opzioni DHCP del VPC al momento della creazione del connettore. Prima di creare un connettore, assicurati di configurare il set di opzioni DHCP del VPC per i requisiti di risoluzione del nome host DNS del connettore.

I connettori creati prima che la funzionalità del nome host Private DNS fosse disponibile in MSK Connect continuano a utilizzare la precedente configurazione di risoluzione DNS senza che sia necessaria alcuna modifica.

Se nel connettore hai bisogno soltanto di una risoluzione dei nomi host DNS risolvibile pubblicamente, per semplificare la configurazione ti consigliamo di utilizzare il VPC predefinito del tuo account quando crei il connettore. Per ulteriori informazioni sul server DNS fornito da Amazon o Risolutore Amazon Route 53, consulta la pagina Amazon DNS Server nella Guida per l'utente di Amazon VPC.

Se devi risolvere nomi host DNS privati, assicurati che le opzioni DHCP del VPC passato durante la creazione del connettore siano impostate correttamente. Per ulteriori informazioni, consulta la pagina Work with DHCP option sets nella Guida per l'utente di Amazon VPC.

Quando configuri un set di opzioni DHCP per la risoluzione dei nomi host DNS privati, assicurati che il connettore possa raggiungere i server DNS personalizzati configurati nel set di opzioni DHCP. In caso contrario, la creazione del connettore avrà esito negativo.

Dopo aver personalizzato il set di opzioni DHCP del VPC, i connettori successivamente creati in tale VPC utilizzano i server DNS specificati nell'insieme di opzioni. Se modifichi il set di opzioni dopo aver creato un connettore, il connettore adotta le impostazioni del nuovo set di opzioni entro un paio di minuti.

Attributi DNS per il VPC

Assicurati di avere configurato correttamente gli attributi DNS del VPC come descritto nelle sezioni DNS attributes in your VPC e DNS hostnames nella Guida per l'utente di Amazon VPC.

Per informazioni sull'uso degli endpoint risolutori in entrata e in uscita per connettere altre reti al tuo VPC e lavorare con il tuo connettore, consulta la pagina Resolving DNS queries between VPCs and your network nella Guida per gli sviluppatori di Amazon Route 53.

Gestione dei guasti

Questa sezione descrive i possibili errori di creazione dei connettori associati alla risoluzione DNS e le operazioni suggerite per risolvere i problemi.

Errore	Operazione suggerita
La creazione del connettore ha esito negativo se una query di risoluzione DNS non riesce o se i server DNS non sono raggiungibili dal connettore.	Puoi vedere gli errori di creazione dei connettori dovuti a query di risoluzione DNS non riuscite nei tuoi CloudWatch log, se hai configurato questi registri per il tuo connettore. Controlla le configurazioni del server DNS e verifica la connettività di rete ai server DNS dal connettore.
Se si modifica la configurazione dei server DNS nel set di opzioni DHCP del VPC mentre un connettore è in esecuzione, le query di risoluzione DNS dal connettore possono avere esito negativo. Se la risoluzione DNS non riesce, alcune attività del connettore possono entrare in uno stato di errore.	Se hai configurato questi registri per il connettore, puoi visualizzare gli errori di creazione dei connettori dovuti a query di risoluzione DNS non riuscite nei tuoi CloudWatch registri. Le attività non riuscite dovrebbero riavviarsi automaticamente per riattivare il connettore. Se ciò non accade, puoi contattare il Supporto

Errore	Operazione suggerita
	per riavviare le attività non riuscite relative al connettore oppure puoi ricreare il connettore.

Registrazione per MSK Connect

MSK Connect è in grado di scrivere log eventi che è possibile utilizzare per eseguire il debug del connettore. Quando si crea un connettore, è possibile specificare nessuna, una o più delle seguenti destinazioni di log:

- Amazon CloudWatch Logs: specificate il gruppo di log a cui desiderate che MSK Connect invii gli eventi di registro del connettore. Per informazioni su come creare un gruppo di log, consulta Create a log group nella CloudWatch Logs User Guide.
- Amazon S3: specifica il bucket S3 a cui desideri che MSK Connect invii i log eventi del connettore. Per informazioni su come creare un bucket S3, consulta la pagina Creating a bucket nella Guida per l'utente di Amazon S3.
- Amazon Data Firehose: specifichi il flusso di distribuzione a cui desideri che MSK Connect invii gli eventi di registro del connettore. Per informazioni su come creare un flusso di distribuzione, consulta Creating an Amazon Data Firehose Delivery stream nella Firehose User Guide.

Per ulteriori informazioni sulla configurazione della registrazione, consulta la pagina Abilitazione della registrazione dai servizi AWS nella Guida per l'utente di Amazon CloudWatch Logs .

MSK Connect emette i seguenti tipi di log eventi:

Livello	Descrizione
`INFO`	Eventi di runtime di interesse all'avvio e all'arresto.
`WARN`	Situazioni di runtime che non sono errori ma sono indesiderate o impreviste.
`FATAL`	Errori gravi che causano una terminazione anticipata.

Livello	Descrizione
ERROR	Condizioni impreviste ed errori di runtime non fatali.

Di seguito è riportato un esempio di evento di registro inviato a CloudWatch Logs:

```
[Worker-0bb8afa0b01391c41] [2021-09-06 16:02:54,151] WARN [Producer
 clientId=producer-1] Connection to node 1 (b-1.my-test-cluster.twwhtj.c2.kafka.us-
east-1.amazonaws.com/INTERNAL_IP) could not be established. Broker may not be
 available. (org.apache.kafka.clients.NetworkClient:782)
```

Impedire la visualizzazione di segreti nei log dei connettori

> **Note**
>
> Se un plug-in non definisce i valori di configurazione sensibili come segreti, tali valori possono apparire nei log dei connettori. Kafka Connect tratta i valori di configurazione non definiti allo stesso modo di qualsiasi altro valore non crittografato.

Se il plug-in definisce una proprietà come segreta, Kafka Connect oscura il valore della proprietà nei log dei connettori. Ad esempio, i seguenti log dei connettori mostrano che se un plug-in definisce `aws.secret.key` come un tipo `PASSWORD`, il suo valore viene sostituito con **`[hidden]`**.

```
   2022-01-11T15:18:55.000+00:00    [Worker-05e6586a48b5f331b] [2022-01-11
15:18:55,150] INFO SecretsManagerConfigProviderConfig values:
   2022-01-11T15:18:55.000+00:00    [Worker-05e6586a48b5f331b] aws.access.key =
my_access_key
   2022-01-11T15:18:55.000+00:00    [Worker-05e6586a48b5f331b] aws.region = us-east-1
   2022-01-11T15:18:55.000+00:00    [Worker-05e6586a48b5f331b] aws.secret.key
= [hidden]
   2022-01-11T15:18:55.000+00:00    [Worker-05e6586a48b5f331b] secret.prefix =
   2022-01-11T15:18:55.000+00:00    [Worker-05e6586a48b5f331b] secret.ttl.ms = 300000
   2022-01-11T15:18:55.000+00:00    [Worker-05e6586a48b5f331b]
(com.github.jcustenborder.kafka.config.aws.SecretsManagerConfigProviderConfig:361)
```

Per evitare che nei file di log dei connettori appaiano dei segreti, gli sviluppatori di plug-in devono utilizzare la costante di enumerazione `ConfigDef.Type.PASSWORD` di Kafka Connect per definire

le proprietà sensibili. Quando una proprietà è di tipo `ConfigDef.Type.PASSWORD`, Kafka Connect esclude il relativo valore dai log dei connettori anche se il valore viene inviato come testo non crittografato.

Monitoraggio di MSK Connect

Il monitoraggio è una parte importante per mantenere l'affidabilità, la disponibilità e le prestazioni di MSK Connect e delle altre AWS soluzioni. Amazon CloudWatch monitora AWS le tue risorse e le applicazioni su cui esegui AWS in tempo reale. Puoi raccogliere i parametri e tenerne traccia, creare pannelli di controllo personalizzati e impostare allarmi per inviare una notifica o intraprendere azioni quando un parametro specificato raggiunge una determinata soglia. Ad esempio, puoi tenere CloudWatch traccia dell'utilizzo della CPU o di altri parametri del connettore, in modo da aumentarne la capacità se necessario. Per ulteriori informazioni, consulta la Amazon CloudWatch User Guide.

La tabella seguente mostra le metriche inviate da MSK CloudWatch Connect all'interno della `ConnectorName` dimensione. MSK Connect fornisce queste metriche per impostazione predefinita e senza costi aggiuntivi. CloudWatch conserva queste metriche per 15 mesi, in modo da poter accedere alle informazioni storiche e avere una prospettiva migliore sulle prestazioni dei connettori. È anche possibile impostare allarmi che controllano determinate soglie e inviare notifiche o intraprendere azioni quando queste soglie vengono raggiunte. Per ulteriori informazioni, consulta la Amazon CloudWatch User Guide.

Parametri di MSK Connect

Nome parametro	Descrizione
`BytesInPerSec`	Il numero di byte ricevuti dal connettore.
`BytesOutPerSec`	Il numero totale di byte distribuiti dal connettor e.
`CpuUtilization`	La percentuale di utilizzo della CPU per sistema e utente.
`ErroredTaskCount`	Il numero di attività che sono state eseguite con errori.
`MemoryUtilization`	La percentuale della memoria totale su un'istanza worker, non solo la memoria heap

Nome parametro	Descrizione
	della macchina virtuale Java (JVM) attualmente in uso. JVM in genere non restituisce la memoria al sistema operativo. Quindi, JVM heap size (MemoryUtilization) di solito inizia con una dimensione minima dell'heap che aumenta in modo incrementale fino a un massimo stabile di circa l'80-90%. L'utilizzo dell'heap JVM potrebbe aumentare o diminuire al variare dell'utilizzo effettivo della memoria da parte del connettore.
`RebalanceCompletedTotal`	Il numero totale di ribilanciamenti completati da questo connettore.
`RebalanceTimeAvg`	Il tempo medio in millisecondi impiegato dal connettore per il ribilanciamento.
`RebalanceTimeMax`	Il tempo massimo in millisecondi impiegato dal connettore per il ribilanciamento.
`RebalanceTimeSinceLast`	Il tempo in millisecondi trascorso dal momento in cui questo connettore ha completato il ribilanciamento più recente.
`RunningTaskCount`	Il numero di attività in esecuzione nel connettore.
`SinkRecordReadRate`	Il numero medio al secondo di record letti dal cluster Apache Kafka o Amazon MSK.
`SinkRecordSendRate`	Il numero medio al secondo di record emessi dalle trasformazioni e inviati alla destinazione. Questo numero non include i record filtrati.
`SourceRecordPollRate`	Il numero medio al secondo di record prodotti o sottoposti a polling.

Nome parametro	Descrizione
`SourceRecordWriteRate`	Il numero medio al secondo di record derivati dalle trasformazioni e scritti sul cluster Apache Kafka o Amazon MSK.
`TaskStartupAttemptsTotal`	Il numero totale di tentativi di avvio di attività eseguiti dal connettore. È possibile utilizzare questo parametro per identificare le anomalie nei tentativi di avvio delle attività.
`TaskStartupSuccessPercentage`	La percentuale media di attività avviate correttamente dal connettore. È possibile utilizzare questo parametro per identificare le anomalie nei tentativi di avvio delle attività.
`WorkerCount`	Il numero minimo di worker in esecuzione nel connettore.

Esempi

Questa sezione include esempi per aiutarti a configurare risorse Amazon MSK Connect come connettori e provider di configurazione di terze parti comuni.

Argomenti

- Connettore sink Amazon S3
- Connettore di origine Debezium con provider di configurazione

Connettore sink Amazon S3

Questo esempio mostra come utilizzare il connettore sink Amazon S3 Confluent e AWS CLI come creare un connettore sink Amazon S3 in MSK Connect.

1. Copia il codice JSON seguente e incollalo in un nuovo file. Sostituisci le stringhe segnaposto con valori che corrispondono alla stringa di connessione dei server di bootstrap del cluster Amazon MSK e agli ID di sottorete e dei gruppi di sicurezza del cluster. Per informazioni su come

configurare un ruolo di esecuzione del servizio, consulta la pagina the section called “Ruoli IAM e policy”.

```
{
    "connectorConfiguration": {
        "connector.class": "io.confluent.connect.s3.S3SinkConnector",
        "s3.region": "us-east-1",
        "format.class": "io.confluent.connect.s3.format.json.JsonFormat",
        "flush.size": "1",
        "schema.compatibility": "NONE",
        "topics": "my-test-topic",
        "tasks.max": "2",
        "partitioner.class":
 "io.confluent.connect.storage.partitioner.DefaultPartitioner",
        "storage.class": "io.confluent.connect.s3.storage.S3Storage",
        "s3.bucket.name": "my-test-bucket"
    },
    "connectorName": "example-S3-sink-connector",
    "kafkaCluster": {
        "apacheKafkaCluster": {
            "bootstrapServers": "<cluster-bootstrap-servers-string>",
            "vpc": {
                "subnets": [
                    "<cluster-subnet-1>",
                    "<cluster-subnet-2>",
                    "<cluster-subnet-3>"
                ],
                "securityGroups": ["<cluster-security-group-id>"]
            }
        }
    },
    "capacity": {
        "provisionedCapacity": {
            "mcuCount": 2,
            "workerCount": 4
        }
    },
    "kafkaConnectVersion": "2.7.1",
    "serviceExecutionRoleArn": "<arn-of-a-role-that-msk-connect-can-assume>",
    "plugins": [
        {
            "customPlugin": {
```

```
                "customPluginArn": "<arn-of-custom-plugin-that-contains-connector-
code>",
                "revision": 1
            }
        }
    ],
    "kafkaClusterEncryptionInTransit": {"encryptionType": "PLAINTEXT"},
    "kafkaClusterClientAuthentication": {"authenticationType": "NONE"}
}
```

2. Esegui il AWS CLI comando seguente nella cartella in cui hai salvato il file JSON nel passaggio precedente.

```
aws kafkaconnect create-connector --cli-input-json file://connector-info.json
```

Di seguito è riportato un esempio dell'output che si ottiene eseguendo correttamente il comando.

```
{
    "ConnectorArn": "arn:aws:kafkaconnect:us-east-1:123450006789:connector/example-
S3-sink-connector/abc12345-abcd-4444-a8b9-123456f513ed-2",
    "ConnectorState": "CREATING",
    "ConnectorName": "example-S3-sink-connector"
}
```

Connettore di origine Debezium con provider di configurazione

Questo esempio mostra come utilizzare il plug-in del connettore Debezium MySQL con un database Amazon Aurora compatibile con MySQL come origine. In questo esempio, abbiamo anche configurato il provider open source AWS Secrets Manager Config Provider per esternalizzare le credenziali del database in AWS Secrets Manager. Per ulteriori informazioni sui provider di configurazione, consulta la pagina Esternalizzazione di informazioni sensibili utilizzando i provider di configurazione.

> **Important**
>
> Il plug-in del connettore Debezium MySQL supporta solo un'attività e non funziona con la modalità di capacità con dimensionamento automatico per Amazon MSK Connect. Dovresti invece utilizzare la modalità di capacità assegnata e impostare il valore `workerCount`

su uno una nella configurazione del connettore. Per ulteriori informazioni sulle modalità di capacità di MSK Connect, consulta la pagina Capacità del connettore.

Prima di iniziare

Il connettore deve essere in grado di accedere a Internet in modo da poter interagire con servizi esterni all'utente Amazon Virtual Private Cloud, ad esempio. AWS Secrets Manager I passaggi descritti in questa sezione consentono di completare le seguenti attività per abilitare l'accesso a Internet.

- Configura una sottorete pubblica che ospita un gateway NAT e indirizza il traffico verso un gateway Internet nel tuo VPC.
- Crea una route predefinita che indirizza il traffico della sottorete privata verso il gateway NAT.

Per ulteriori informazioni, consulta Abilitazione dell'accesso a Internet per Amazon MSK Connect.

Prerequisiti

Prima di abilitare l'accesso a Internet, devi disporre dei seguenti elementi:

- L'ID del Amazon Virtual Private Cloud (VPC) associato al cluster. Ad esempio, vpc-123456ab.
- Gli ID delle sottoreti private nel VPC. Ad esempio, subnet-a1b2c3de, subnet-f4g5h6ij e così via. Il connettore deve essere configurato con sottoreti private.

Abilitazione dell'accesso a Internet per il connettore

1. Apri la Amazon Virtual Private Cloud console all'indirizzo https://console.aws.amazon.com/vpc/.
2. Crea una sottorete pubblica con un nome descrittivo per il gateway NAT e prendi nota dell'ID della sottorete. Per istruzioni dettagliate, consulta la pagina Create a subnet in your VPC.
3. Crea un gateway Internet in modo che il VPC possa comunicare con Internet e prendi nota dell'ID del gateway. Collega il gateway Internet al VPC. Per istruzioni, consulta la pagina Create and attach an internet gateway.
4. Fornisci un gateway NAT pubblico in modo che gli host delle tue sottoreti private possano raggiungere la tua sottorete pubblica. Quando crei il gateway NAT, seleziona la sottorete pubblica creata in precedenza. Per istruzioni, consulta Creazione di un gateway NAT.

5. Configura le tabelle di routing. Per completare questa configurazione, occorrono in totale due tabelle di routing. Dovresti già disporre di una tabella di routing principale creata in automatico al momento della creazione del VPC. In questo passaggio creerai una tabella di routing aggiuntiva per la sottorete pubblica.

   a. Utilizza le seguenti impostazioni per modificare la tabella di routing principale del tuo VPC in modo che le sottoreti private instradino il traffico verso il tuo gateway NAT. Per le istruzioni, consulta la pagina Utilizzo delle tabelle di routing nella Guida per l'utente di Amazon Virtual Private Cloud.

      Tabella di routing MSKC privata

Proprietà	Valore
Name tag (Tag nome)	Ti consigliamo di assegnare a questa tabella di routing un nome descrittivo per facilitarne l'identificazione. Ad esempio, MSKC privata.
Sottoreti associate	Le tue sottoreti private
Un percorso per consentire l'accesso a Internet per MSK Connect	• Destinazione: 0.0.0.0/0 • Obiettivo: l'ID del gateway NAT. Ad esempio, nat-12a345bc6789efg1h.
Un percorso per tutto il traffico locale	• Destinazione: 10.0.0.0/16. Questo valore può variare a seconda del blocco CIDR del tuo VPC. • Obiettivo: locale

   b. Segui le istruzioni riportate nella pagina Creazione di una tabella di routing personalizzata per creare una tabella di routing per la sottorete pubblica. Quando crei la tabella, inserisci un nome descrittivo nel campo Tag nome per identificare a quale sottorete è associata la tabella. Ad esempio, MSKC pubblica.

   c. Configura la tua tabella di routing MSKC pubblica utilizzando le seguenti impostazioni.

Proprietà	Valore
Name tag (Tag nome)	MSKC pubblica o un altro nome descrittivo a scelta
Sottoreti associate	La tua sottorete pubblica con gateway NAT
Un percorso per consentire l'accesso a Internet per MSK Connect	• Destinazione: 0.0.0.0/0 • Obiettivo: l'ID del gateway Internet. Ad esempio, igw-1a234bc5.
Un percorso per tutto il traffico locale	• Destinazione: 10.0.0.0/16. Questo valore può variare a seconda del blocco CIDR del tuo VPC. • Obiettivo: locale

Ora che hai abilitato l'accesso a Internet per Amazon MSK Connect, puoi creare un connettore.

Creazione di un connettore di origine Debezium

1. Creazione di un plug-in personalizzato

   a. Scarica il plug-in del connettore MySQL per l'ultima versione stabile dal sito Debezium. Prendi nota della versione di rilascio di Debezium che scarichi (versione 2.x o la vecchia serie 1.x). Più avanti in questa procedura, creerai un connettore basato sulla tua versione di Debezium.

   b. Scarica ed estrai AWS Secrets Manager Config Provider.

   c. Colloca i seguenti archivi nella stessa directory:

      • La cartella `debezium-connector-mysql`

      • La cartella `jcusten-border-kafka-config-provider-aws-0.1.1`

   d. Comprimi la directory che hai creato nel passaggio precedente in un file ZIP, quindi carica il file ZIP in un bucket S3. Per istruzioni, consulta la pagina Uploading objects in Amazon S3 nella Guida per l'utente di Amazon S3.

e. Copia il codice JSON seguente e incollalo in un file. Ad esempio, `debezium-source-custom-plugin.json`. Sostituisci *< example-custom-plugin-name >* con il nome che vuoi che abbia il plugin, *< arn-of-your-s 3-bucket>* con l'ARN del bucket S3 in cui hai caricato il file ZIP e *<file-key-of-ZIP-object>* con la chiave del file dell'oggetto ZIP che hai caricato su S3.

```
{
    "name": "<example-custom-plugin-name>",
    "contentType": "ZIP",
    "location": {
        "s3Location": {
            "bucketArn": "<arn-of-your-s3-bucket>",
            "fileKey": "<file-key-of-ZIP-object>"
        }
    }
}
```

f. Esegui il seguente AWS CLI comando dalla cartella in cui hai salvato il file JSON per creare un plugin.

```
aws kafkaconnect create-custom-plugin --cli-input-json file://<debezium-source-custom-plugin.json>
```

Verrà visualizzato un output simile al seguente.

```
{
    "CustomPluginArn": "arn:aws:kafkaconnect:us-east-1:012345678901:custom-plugin/example-custom-plugin-name/abcd1234-a0b0-1234-c1-12345678abcd-1",
    "CustomPluginState": "CREATING",
    "Name": "example-custom-plugin-name",
    "Revision": 1
}
```

g. Esegui il comando seguente per verificare lo stato del plug-in. Lo stato del cluster dovrebbe passare da `CREATING` a `ACTIVE`. Sostituisci il segnaposto ARN con l'ARN ottenuto nell'output del comando precedente.

```
aws kafkaconnect describe-custom-plugin --custom-plugin-arn "<arn-of-your-custom-plugin>"
```

2. Configura AWS Secrets Manager e crea un segreto per le credenziali del tuo database

   a. Apri la console di Secrets Manager all'indirizzo https://console.aws.amazon.com/secretsmanager/.

   b. Crea un nuovo segreto per archiviare le credenziali di accesso al database. Per le istruzioni, consulta la pagina Create a secret nella Guida per l'utente di AWS Secrets Manager.

   c. Copia l'ARN del segreto.

   d. Aggiungi le autorizzazioni di Secrets Manager dalla seguente policy di esempio al tuo Ruolo di esecuzione del servizio. Sostituisci *<arn:aws:secretsmanager:us-east-1:123456789000:secret*: -1234> con l'ARN del tuo segreto. MySecret

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "secretsmanager:GetResourcePolicy",
        "secretsmanager:GetSecretValue",
        "secretsmanager:DescribeSecret",
        "secretsmanager:ListSecretVersionIds"
      ],
      "Resource": [
        "<arn:aws:secretsmanager:us-east-1:123456789000:secret:MySecret-1234>"
      ]
    }
  ]
}
```

   Per istruzioni sull'aggiunta di autorizzazioni IAM, consulta la pagina Adding and removing IAM identity permissions nella Guida per l'utente di IAM.

3. Creazione di una configurazione del worker personalizzata con informazioni sul proprio provider di configurazione

   a. Copia le seguenti proprietà di configurazione del worker in un file, sostituendo le stringhe segnaposto con valori che corrispondono al tuo scenario. Per ulteriori informazioni sulle proprietà di configurazione per il provider di configurazione di AWS Secrets Manager Config, SecretsManagerConfigProviderconsultate la documentazione del plugin.

```
key.converter=<org.apache.kafka.connect.storage.StringConverter>
value.converter=<org.apache.kafka.connect.storage.StringConverter>
config.providers.secretManager.class=com.github.jcustenborder.kafka.config.aws.SecretsM
config.providers=secretManager
config.providers.secretManager.param.aws.region=<us-east-1>
```

b. Esegui il AWS CLI comando seguente per creare la tua configurazione di lavoro personalizzata.

Sostituisci i valori seguenti:

- *< my-worker-config-name >* - un nome descrittivo per la configurazione personalizzata del lavoratore
- *< encoded-properties-file-content -string>* - una versione con codifica base64 delle proprietà di testo in chiaro copiate nel passaggio precedente

```
aws kafkaconnect create-worker-configuration --name <my-worker-config-name> --
properties-file-content <encoded-properties-file-content-string>
```

4. Creazione di un connettore

a. Copia il codice JSON seguente, che corrisponde alla tua versione di Debezium (2.x o 1.x), e incollalo in un nuovo file. Sostituisci le stringhe *<placeholder>* con valori che corrispondono al tuo scenario. Per informazioni su come configurare un ruolo di esecuzione del servizio, consulta la pagina the section called “Ruoli IAM e policy”.

Nota che la configurazione utilizza variabili come `${secretManager:MySecret-1234:dbusername}` anziché testo non crittografato per specificare le credenziali del database. Sostituisci *MySecret-1234* con il nome del tuo segreto, poi includi il nome della chiave che desideri recuperare. È inoltre necessario sostituire *<arn-of-config-provider-worker-configuration>* con l'ARN della configurazione del worker personalizzata.

Debezium 2.x

Per le versioni di Debezium 2.x, copia il codice JSON seguente e incollalo in un nuovo file. Sostituisci le stringhe *<placeholder>* con valori che corrispondono al tuo scenario.

```
{
 "connectorConfiguration": {
  "connector.class": "io.debezium.connector.mysql.MySqlConnector",
  "tasks.max": "1",
  "database.hostname": "<aurora-database-writer-instance-endpoint>",
  "database.port": "3306",
  "database.user": "<${secretManager:MySecret-1234:dbusername}>",
  "database.password": "<${secretManager:MySecret-1234:dbpassword}>",
  "database.server.id": "123456",
  "database.include.list": "<list-of-databases-hosted-by-specified-server>",
  "topic.prefix": "<logical-name-of-database-server>",
  "schema.history.internal.kafka.topic": "<kafka-topic-used-by-debezium-to-
track-schema-changes>",
  "schema.history.internal.kafka.bootstrap.servers": "<cluster-bootstrap-
servers-string>",
  "schema.history.internal.consumer.security.protocol": "SASL_SSL",
  "schema.history.internal.consumer.sasl.mechanism": "AWS_MSK_IAM",
  "schema.history.internal.consumer.sasl.jaas.config":
 "software.amazon.msk.auth.iam.IAMLoginModule required;",
  "schema.history.internal.consumer.sasl.client.callback.handler.class":
 "software.amazon.msk.auth.iam.IAMClientCallbackHandler",
  "schema.history.internal.producer.security.protocol": "SASL_SSL",
  "schema.history.internal.producer.sasl.mechanism": "AWS_MSK_IAM",
  "schema.history.internal.producer.sasl.jaas.config":
 "software.amazon.msk.auth.iam.IAMLoginModule required;",
  "schema.history.internal.producer.sasl.client.callback.handler.class":
 "software.amazon.msk.auth.iam.IAMClientCallbackHandler",
  "include.schema.changes": "true"
 },
 "connectorName": "example-Debezium-source-connector",
 "kafkaCluster": {
  "apacheKafkaCluster": {
   "bootstrapServers": "<cluster-bootstrap-servers-string>",
   "vpc": {
    "subnets": [
     "<cluster-subnet-1>",
     "<cluster-subnet-2>",
     "<cluster-subnet-3>"
    ],
    "securityGroups": ["<id-of-cluster-security-group>"]
   }
  }
 },
```

```
 "capacity": {
  "provisionedCapacity": {
   "mcuCount": 2,
   "workerCount": 1
  }
 },
 "kafkaConnectVersion": "2.7.1",
 "serviceExecutionRoleArn": "<arn-of-service-execution-role-that-msk-
connect-can-assume>",
 "plugins": [{
  "customPlugin": {
   "customPluginArn": "<arn-of-msk-connect-plugin-that-contains-connector-
code>",
   "revision": 1
  }
 }],
 "kafkaClusterEncryptionInTransit": {
  "encryptionType": "TLS"
 },
 "kafkaClusterClientAuthentication": {
  "authenticationType": "IAM"
 },
 "workerConfiguration": {
  "workerConfigurationArn": "<arn-of-config-provider-worker-configuration>",
  "revision": 1
 }
}
```

Debezium 1.x

Per le versioni di Debezium 1.x, copia il codice JSON seguente e incollalo in un nuovo file. Sostituisci le stringhe *<placeholder>* con valori che corrispondono al tuo scenario.

```
{
 "connectorConfiguration": {
  "connector.class": "io.debezium.connector.mysql.MySqlConnector",
  "tasks.max": "1",
  "database.hostname": "<aurora-database-writer-instance-endpoint>",
  "database.port": "3306",
  "database.user": "<${secretManager:MySecret-1234:dbusername}>",
  "database.password": "<${secretManager:MySecret-1234:dbpassword}>",
  "database.server.id": "123456",
  "database.server.name": "<logical-name-of-database-server>",
```

```
  "database.include.list": "<list-of-databases-hosted-by-specified-server>",
  "database.history.kafka.topic": "<kafka-topic-used-by-debezium-to-track-
schema-changes>",
  "database.history.kafka.bootstrap.servers": "<cluster-bootstrap-servers-
string>",
  "database.history.consumer.security.protocol": "SASL_SSL",
  "database.history.consumer.sasl.mechanism": "AWS_MSK_IAM",
  "database.history.consumer.sasl.jaas.config":
 "software.amazon.msk.auth.iam.IAMLoginModule required;",
  "database.history.consumer.sasl.client.callback.handler.class":
 "software.amazon.msk.auth.iam.IAMClientCallbackHandler",
  "database.history.producer.security.protocol": "SASL_SSL",
  "database.history.producer.sasl.mechanism": "AWS_MSK_IAM",
  "database.history.producer.sasl.jaas.config":
 "software.amazon.msk.auth.iam.IAMLoginModule required;",
  "database.history.producer.sasl.client.callback.handler.class":
 "software.amazon.msk.auth.iam.IAMClientCallbackHandler",
  "include.schema.changes": "true"
 },
 "connectorName": "example-Debezium-source-connector",
 "kafkaCluster": {
  "apacheKafkaCluster": {
   "bootstrapServers": "<cluster-bootstrap-servers-string>",
   "vpc": {
    "subnets": [
     "<cluster-subnet-1>",
     "<cluster-subnet-2>",
     "<cluster-subnet-3>"
    ],
    "securityGroups": ["<id-of-cluster-security-group>"]
   }
  }
 },
 "capacity": {
  "provisionedCapacity": {
   "mcuCount": 2,
   "workerCount": 1
  }
 },
 "kafkaConnectVersion": "2.7.1",
 "serviceExecutionRoleArn": "<arn-of-service-execution-role-that-msk-
connect-can-assume>",
 "plugins": [{
  "customPlugin": {
```

```
   "customPluginArn": "<arn-of-msk-connect-plugin-that-contains-connector-
code>",
   "revision": 1
  }
 }],
 "kafkaClusterEncryptionInTransit": {
  "encryptionType": "TLS"
 },
 "kafkaClusterClientAuthentication": {
  "authenticationType": "IAM"
 },
 "workerConfiguration": {
  "workerConfigurationArn": "<arn-of-config-provider-worker-configuration>",
  "revision": 1
 }
}
```

b. Esegui il AWS CLI comando seguente nella cartella in cui hai salvato il file JSON nel passaggio precedente.

```
aws kafkaconnect create-connector --cli-input-json file://connector-info.json
```

Di seguito è riportato un esempio dell'output che si ottiene eseguendo correttamente il comando.

```
{
    "ConnectorArn": "arn:aws:kafkaconnect:us-east-1:123450006789:connector/
example-Debezium-source-connector/abc12345-abcd-4444-a8b9-123456f513ed-2",
    "ConnectorState": "CREATING",
    "ConnectorName": "example-Debezium-source-connector"
}
```

Per un esempio di connettore Debezium con i passaggi dettagliati, consulta la pagina Introducing Amazon MSK Connect - Stream Data to and from Your Apache Kafka Clusters Using Managed Connectors.

Best practice

Utilizza queste informazioni come riferimento per trovare rapidamente le raccomandazioni per ottimizzare le prestazioni di Amazon MSK Connect.

Argomenti

Connessione dai connettori

Le seguenti best practice possono migliorare le prestazioni della connettività ad Amazon MSK Connect.

Evitare di sovrapporre gli IP per il peering o il gateway di transito di Amazon VPC

Se utilizzi il peering o il gateway di transito di Amazon VPC con Amazon MSK Connect, evita di configurare il connettore per raggiungere le risorse VPC in peering con IP negli intervalli CIDR:

- "10.99.0.0/16"
- "192.168.0.0/16"
- "172.21.0.0/16"

Guida alla migrazione di Amazon MSK Connect

Questa sezione descrive come migrare l'applicazione del connettore Apache Kafka su Amazon Managed Streaming for Apache Kafka Connect (Amazon MSK Connect).

Argomenti

Vantaggi dell'utilizzo di Amazon MSK Connect

Apache Kafka è una delle piattaforme di streaming open source più utilizzate per l'acquisizione e l'elaborazione di flussi di dati in tempo reale. Con Apache Kafka, puoi disaccoppiare e scalare in modo indipendente le tue applicazioni che producono e consumano dati.

Kafka Connect è un componente importante per la creazione e l'esecuzione di applicazioni di streaming con Apache Kafka. Kafka Connect offre un modo standardizzato per lo spostamento dei dati tra Kafka e sistemi esterni. Kafka Connect è altamente scalabile e può gestire grandi volumi

di dati Kafka Connect fornisce un potente set di operazioni e strumenti API per la configurazione, l'implementazione e il monitoraggio dei connettori che spostano i dati tra argomenti Kafka e sistemi esterni. Puoi utilizzare questi strumenti per personalizzare ed estendere le funzionalità di Kafka Connect per soddisfare le esigenze specifiche della tua applicazione di streaming.

Potresti incontrare delle difficoltà quando gestisci i cluster Apache Kafka Connect da soli o quando cerchi di migrare applicazioni open source Apache Kafka Connect verso. AWS Queste sfide includono il tempo necessario per configurare l'infrastruttura e implementare le applicazioni, gli ostacoli tecnici alla configurazione dei cluster Apache Kafka Connect autogestiti e il sovraccarico operativo amministrativo.

Per affrontare queste sfide, ti consigliamo di utilizzare Amazon Managed Streaming for Apache Kafka Connect (Amazon MSK Connect) per migrare le tue applicazioni open source Apache Kafka Connect verso. AWS Amazon MSK Connect semplifica l'utilizzo di Kafka Connect per lo streaming di dati da e verso cluster Apache Kafka e sistemi esterni, come database, indici di ricerca e file system.

Ecco alcuni dei vantaggi della migrazione ad Amazon MSK Connect:

- Eliminazione del sovraccarico operativo: Amazon MSK Connect elimina il carico operativo associato all'applicazione di patch, al provisioning e al ridimensionamento dei cluster Apache Kafka Connect. Amazon MSK Connect monitora continuamente lo stato dei cluster Connect e automatizza l'applicazione di patch e aggiornamenti di versione senza causare interruzioni ai carichi di lavoro.
- Riavvio automatico delle attività di Connect: Amazon MSK Connect può ripristinare automaticamente le attività non riuscite per ridurre le interruzioni della produzione. Gli errori delle attività possono essere causati da errori temporanei, come il superamento del limite di connessione TCP per Kafka, e il ribilanciamento delle attività quando nuovi lavoratori si uniscono al gruppo di consumatori per i connettori sink.
- Scalabilità orizzontale e verticale automatica: Amazon MSK Connect consente all'applicazione del connettore di scalare automaticamente per supportare throughput più elevati. Amazon MSK Connect gestisce la scalabilità per te. È sufficiente specificare il numero di lavoratori nel gruppo di auto scaling e le soglie di utilizzo. Puoi utilizzare l'operazione dell'`UpdateConnector`API Amazon MSK Connect per scalare verticalmente verso l'alto o verso il basso le vCPU tra 1 e 8 vCPU per supportare un throughput variabile.
- Connettività di rete privata: Amazon MSK Connect si connette privatamente ai sistemi di origine e sink utilizzando nomi AWS PrivateLink DNS privati.

Migrazione ad Amazon MSK Connect

Questa sezione descrive brevemente gli argomenti di gestione dello stato utilizzati da Kafka Connect e Amazon MSK Connect. Questa sezione descrive anche le procedure per la migrazione dei connettori source e sink.

Argomenti

Argomenti interni utilizzati da Kafka Connect

Un'applicazione Apache Kafka Connect in esecuzione in modalità distribuita memorizza il proprio stato utilizzando argomenti interni nel cluster Kafka e l'appartenenza ai gruppi. I seguenti sono i valori di configurazione che corrispondono agli argomenti interni utilizzati per le applicazioni Kafka Connect:

- Argomento di configurazione, specificato tramite `config.storage.topic`

  Nell'argomento di configurazione, Kafka Connect memorizza la configurazione di tutti i connettori e le attività avviate dagli utenti. Ogni volta che gli utenti aggiornano la configurazione di un connettore o quando un connettore richiede una riconfigurazione (ad esempio, il connettore rileva che può avviare più attività), viene emesso un record su questo argomento. Questo argomento è abilitato alla compattazione, quindi mantiene sempre l'ultimo stato per ogni entità.
- Argomento Offsets, specificato tramite `offset.storage.topic`

  Nell'argomento offset, Kafka Connect memorizza gli offset dei connettori sorgente. Come l'argomento sulla configurazione, l'argomento offset è abilitato alla compattazione. Questo argomento viene utilizzato per scrivere le posizioni di origine solo per i connettori di origine che producono dati a Kafka da sistemi esterni. I connettori Sink, che leggono i dati da Kafka e li inviano a sistemi esterni, memorizzano i dati di consumo utilizzando i normali gruppi di consumatori Kafka.
- Argomento relativo allo stato, specificato tramite `status.storage.topic`

  Nell'argomento relativo allo stato, Kafka Connect memorizza lo stato corrente dei connettori e delle attività. Questo argomento viene utilizzato come punto centrale per i dati richiesti dagli utenti dell'API REST. Questo argomento consente agli utenti di interrogare qualsiasi worker e ottenere

comunque lo stato di tutti i plugin in esecuzione. Come gli argomenti di configurazione e offset, anche l'argomento status è abilitato alla compattazione.

Oltre a questi argomenti, Kafka Connect fa ampio uso dell'API di appartenenza ai gruppi di Kafka. I gruppi prendono il nome dal nome del connettore. Ad esempio, per un connettore denominato file-sink, il gruppo viene denominato. connect-file-sink Ogni consumatore del gruppo fornisce i record relativi a una singola attività. Questi gruppi e i relativi offset possono essere recuperati utilizzando i normali strumenti dei gruppi di consumatori, come. `Kafka-consumer-group.sh` Per ogni connettore sink, il runtime Connect esegue un normale gruppo di consumatori che estrae i record da Kafka.

Gestione dello stato delle applicazioni Amazon MSK Connect

Per impostazione predefinita, Amazon MSK Connect crea tre argomenti separati nel cluster Kafka per ogni connettore Amazon MSK per memorizzare la configurazione, l'offset e lo stato del connettore. I nomi degli argomenti predefiniti sono strutturati come segue:

- *__msk_connect_configs_ nome-connettore _ id-connettore*
- *__msk_connect_status_ nome-connettore _ id-connettore*
- *__msk_connect_offsets_ nome-connettore _ id-connettore*

> **Note**
>
> Per fornire la continuità dell'offset tra i connettori di origine, puoi utilizzare un argomento di memorizzazione dell'offset a tua scelta, anziché l'argomento predefinito. La definizione di un argomento di archiviazione degli offset consente di eseguire attività come la creazione di un connettore di origine che riprenda la lettura dall'ultimo offset di un connettore precedente. Per specificare un argomento di storage offset, fornisci un valore per la offset.storage.topicproprietà nella configurazione del worker Amazon MSK Connect prima di creare il connettore.

Migrazione dei connettori di origine ad Amazon MSK Connect

I connettori di origine sono applicazioni Apache Kafka Connect che importano record da sistemi esterni in Kafka. Questa sezione descrive il processo di migrazione delle applicazioni del connettore

di origine Apache Kafka Connect che eseguono cluster Kafka Connect locali o autogestiti in esecuzione su Amazon MSK Connect. AWS

L'applicazione Kafka Connect source connector memorizza gli offset in un argomento denominato con il valore impostato per la proprietà config. `offset.storage.topic` Di seguito sono riportati alcuni esempi di messaggi di offset per un connettore JDBC che esegue due attività che importano dati da due tabelle diverse denominate e. `movies` `shows` La riga più recente importata dai film da tavolo ha un ID primario di. `18343` La riga più recente importata dalla tabella shows ha un ID primario di`732`.

```
["jdbcsource",{"protocol":"1","table":"sample.movies"}] {"incrementing":18343}
["jdbcsource",{"protocol":"1","table":"sample.shows"}] {"incrementing":732}
```

Per migrare i connettori di origine su Amazon MSK Connect, procedi come segue:

1. Crea un plug-in personalizzato Amazon MSK Connect estraendo le librerie di connettori dal tuo cluster Kafka Connect locale o autogestito.
2. Crea le proprietà dei lavoratori Amazon MSK Connect e imposta le proprietà `key.converter` e `offset.storage.topic` gli stessi valori impostati per il connettore Kafka in esecuzione nel tuo cluster Kafka Connect esistente. `value.converter`
3. Metti in pausa l'applicazione del connettore sul cluster esistente effettuando una `PUT /connectors/`*`connector-name`*`/pause` richiesta sul cluster Kafka Connect esistente.
4. Assicurati che tutte le attività dell'applicazione del connettore siano completamente interrotte. È possibile interrompere le attività effettuando una `GET /connectors/`*`connector-name`*`/status` richiesta sul cluster Kafka Connect esistente o consumando i messaggi dal nome dell'argomento impostato per la proprietà. `status.storage.topic`
5. Ottieni la configurazione del connettore dal cluster esistente. È possibile ottenere la configurazione del connettore effettuando una `GET /connectors/`*`connector-name`*`/config/` richiesta sul cluster esistente o utilizzando i messaggi dal nome dell'argomento impostato per la proprietà`config.storage.topic`.
6. Crea un nuovo Amazon MSK Connector con lo stesso nome di un cluster esistente. Crea questo connettore utilizzando il plug-in personalizzato del connettore che hai creato nel passaggio 1, le proprietà del worker che hai creato nel passaggio 2 e la configurazione del connettore che hai estratto nel passaggio 5.
7. Quando lo stato di Amazon MSK Connector è `active` impostato su, visualizza i log per verificare che il connettore abbia iniziato a importare dati dal sistema di origine.

8. Elimina il connettore nel cluster esistente effettuando una richiesta. `DELETE /connectors/connector-name`

Migrazione dei connettori sink ad Amazon MSK Connect

I connettori Sink sono applicazioni Apache Kafka Connect che esportano dati da Kafka a sistemi esterni. Questa sezione descrive il processo di migrazione delle applicazioni sink connector di Apache Kafka Connect che eseguono cluster Kafka Connect locali o autogestiti in esecuzione su Amazon MSK Connect. AWS

I connettori sink Kafka Connect utilizzano l'API Kafka Group Membership e memorizzano gli offset negli stessi `__consumer_offset` argomenti di una tipica applicazione consumer. Questo comportamento semplifica la migrazione del connettore sink da un cluster autogestito ad Amazon MSK Connect.

Per migrare i connettori sink su Amazon MSK Connect, procedi come segue:

1. Crea un plug-in personalizzato Amazon MSK Connect estraendo le librerie di connettori dal tuo cluster Kafka Connect locale o autogestito.
2. Crea le proprietà dei lavoratori Amazon MSK Connect e imposta le proprietà `key.converter` e `value.converter` gli stessi valori impostati per il connettore Kafka in esecuzione nel tuo cluster Kafka Connect esistente.
3. Metti in pausa l'applicazione del connettore sul cluster esistente effettuando una `PUT /connectors/connector-name/pause` richiesta sul cluster Kafka Connect esistente.
4. Assicurati che tutte le attività dell'applicazione del connettore siano completamente interrotte. È possibile interrompere le attività effettuando una `GET /connectors/connector-name/status` richiesta sul cluster Kafka Connect esistente o consumando i messaggi dal nome dell'argomento impostato per la proprietà. `status.storage.topic`
5. Ottieni la configurazione del connettore dal cluster esistente. È possibile ottenere la configurazione del connettore effettuando una `GET /connectors/connector-name/config` richiesta sul cluster esistente o utilizzando i messaggi dal nome dell'argomento impostato per la proprietà`config.storage.topic`.
6. Crea un nuovo Amazon MSK Connector con lo stesso nome del cluster esistente. Crea questo connettore utilizzando il plug-in personalizzato del connettore che hai creato nel passaggio 1, le proprietà del worker che hai creato nel passaggio 2 e la configurazione del connettore che hai estratto nel passaggio 5.

7. Quando lo stato di Amazon MSK Connector è `active` impostato su, visualizza i log per verificare che il connettore abbia iniziato a importare dati dal sistema di origine.
8. Elimina il connettore nel cluster esistente effettuando una richiesta. `DELETE /connectors/`*`connector-name`*

Risoluzione dei problemi relativi ad Amazon MSK Connect

Le seguenti informazioni agevolano la risoluzione dei problemi che si potrebbero verificare durante l'utilizzo di MSK Connect. Puoi anche pubblicare il problema nel AWS re:Post.

Il connettore non è in grado di accedere alle risorse ospitate sulla rete Internet pubblica

Consulta la sezione Abilitazione dell'accesso a Internet per Amazon MSK Connect.

Il numero di attività in esecuzione del connettore non corrisponde al numero di attività specificato in tasks.max

Ecco alcuni motivi per cui un connettore può utilizzare un numero inferiore di attività rispetto alla configurazione tasks.max specificata:

- Alcune implementazioni di connettori limitano il numero di attività che è possibile utilizzare. Ad esempio, il connettore Debezium per MySQL è limitato all'utilizzo di una singola attività.
- Quando si utilizza la modalità di capacità con dimensionamento automatico, Amazon MSK Connect sovrascrive la proprietà tasks.max del connettore con un valore proporzionale al numero di worker in esecuzione nel connettore e al numero di MCU per worker.
- Per i connettori sink, il livello di parallelismo (numero di attività) non può essere superiore al numero di partizioni di argomento. Sebbene sia possibile impostare tasks.max su un valore maggiore, una singola partizione non viene mai elaborata da più di una singola attività alla volta.
- In Kafka Connect 2.7.x, l'assegnatore di partizioni dei consumatori predefinito è `RangeAssignor`. Il comportamento di questo assegnatore è quello di assegnare la prima partizione di ogni argomento a un singolo consumatore, la seconda partizione di ogni argomento a un singolo consumatore ecc. Ciò significa che il numero massimo di attività attive utilizzate da un connettore sink tramite `RangeAssignor` è uguale al numero massimo di partizioni utilizzate in ogni singolo argomento utilizzato. Se ciò non funziona per il tuo caso d'uso, dovresti creare una configurazione del worker in cui la proprietà `consumer.partition.assignment.strategy` sia impostata su un assegnatore di partizioni dei consumatori più adatto. Vedi Interfaccia Kafka 2.7 ConsumerPartitionAssignor: tutte le classi di implementazione conosciute.

Replicatore MSK

Cos'è il replicatore Amazon MSK?

Amazon MSK Replicator è una funzionalità di Amazon MSK che consente di replicare in modo affidabile i dati tra cluster Amazon MSK in regioni diverse o uguali. AWS Con il replicatore MSK, è possibile creare facilmente applicazioni di streaming resilienti a livello regionale per una maggiore disponibilità e continuità aziendale. Il replicatore MSK fornisce la replica asincrona automatica tra i cluster MSK, eliminando la necessità di scrivere codice personalizzato, gestire l'infrastruttura o configurare reti tra regioni.

Il replicatore MSK dimensiona automaticamente le risorse sottostanti in modo da poter replicare i dati on demand senza dover monitorare o dimensionare la capacità. Il replicatore MSK replica anche i metadati Kafka necessari, tra cui configurazioni degli argomenti, liste di controllo degli accessi (ACL) e offset dei gruppi di consumatori. Se si verifica un evento imprevisto in una regione, puoi eseguire il failover nell'altra regione e riprendere l'elaborazione senza interruzioni. AWS

Il replicatore MSK supporta sia la replica tra regioni (CRR) sia la replica nella stessa regione (SRR). Nella replica tra regioni, i cluster MSK di origine e di destinazione si trovano in regioni diverse. AWS Nella replica nella stessa regione, i cluster MSK di origine e di destinazione si trovano nella stessa regione. AWS È necessario creare cluster MSK di origine e di destinazione prima di poterli utilizzare con il replicatore MSK.

> **Note**
>
> MSK Replicator supporta le seguenti AWS regioni: Stati Uniti orientali (us-east-1, Virginia settentrionale); Stati Uniti orientali (us-east-2, Ohio); Stati Uniti occidentali (us-west-2, Oregon); Europa (eu-west-1, Irlanda); Europa (eu-central-1, Francoforte); Asia Pacifico (ap-southeast-1, Singapore)); Asia-Pacifico (ap-southeast-2, Sydney), Europa (eu-north-1, Stoccolma), Asia Pacifico (ap-south-1, Mumbai), Europa (eu-west-3, Parigi), Sud America (sa-east-1, San Paolo), Asia Pacifico (ap-northeast-2, Sea-1 Oul), Europa (eu-west-2, Londra), Asia Pacifico (ap-northeast-1, Tokyo), Stati Uniti occidentali (us-west-1, California settentrionale), Canada (ca-central-1, centrale).

Ecco alcuni usi comuni del replicatore Amazon MSK.

- Crea applicazioni di streaming tra regioni: crea applicazioni di streaming ad alta disponibilità e tolleranti ai guasti per conseguire una maggiore resilienza senza necessità di configurare soluzioni personalizzate.
- Fornisci un accesso ai dati a bassa latenza: offri un accesso ai dati con latenza inferiore ai consumatori ubicati in aree geografiche diverse.
- Distribuisci i dati ai tuoi partner: copia i dati da un cluster Apache Kafka in più cluster Apache Kafka in modo che diversi team/partner abbiano le proprie copie dei dati.
- Aggrega i dati per l'analisi: copia i dati da più cluster Apache Kafka in un unico cluster per generare facilmente approfondimenti su dati aggregati in tempo reale.
- Scrivi localmente, accedi ai tuoi dati a livello globale: configura la replica multiattiva per propagare automaticamente le scritture eseguite in una AWS regione ad altre regioni per fornire dati a latenza e costi inferiori.

Come funziona il replicatore Amazon MSK

Per iniziare a utilizzare MSK Replicator, è necessario creare un nuovo replicatore nella regione del cluster di destinazione. AWS MSK Replicator copia automaticamente tutti i dati dal cluster nella AWS regione primaria denominata origine nel cluster nella regione di destinazione denominata destinazione. I cluster di origine e di destinazione possono trovarsi nella stessa regione o in regioni diverse. AWS Se il cluster di destinazione non esiste ancora, devi crearlo.

Quando si crea un replicatore, MSK Replicator distribuisce tutte le risorse necessarie nella AWS regione del cluster di destinazione per ottimizzare la latenza di replica dei dati. La latenza di replica varia in base a molti fattori, tra cui la distanza di rete tra le AWS regioni dei cluster MSK, la capacità di throughput dei cluster di origine e di destinazione e il numero di partizioni sui cluster di origine e di destinazione. Il replicatore MSK dimensiona automaticamente le risorse sottostanti in modo da poter replicare i dati on demand senza dover monitorare o dimensionare la capacità.

Replica dei dati

Per impostazione predefinita, MSK Replicator copia tutti i dati in modo asincrono dall'ultimo offset nelle partizioni tematiche del cluster di origine nel cluster di destinazione. Se l'impostazione «Rileva e copia nuovi argomenti» è attivata, MSK Replicator rileva e copia automaticamente nuovi argomenti o partizioni di argomenti nel cluster di destinazione. Tuttavia, il Replicator potrebbe impiegare fino a 30 secondi per rilevare e creare nuovi argomenti o partizioni di argomenti nel cluster di destinazione. Tutti i messaggi inviati all'argomento di origine prima della creazione dell'argomento nel cluster di

destinazione non verranno replicati. In alternativa, è possibile configurare il Replicator durante la creazione per avviare la replica dal primo offset nelle partizioni degli argomenti del cluster di origine se si desidera replicare i messaggi esistenti sui propri argomenti nel cluster di destinazione.

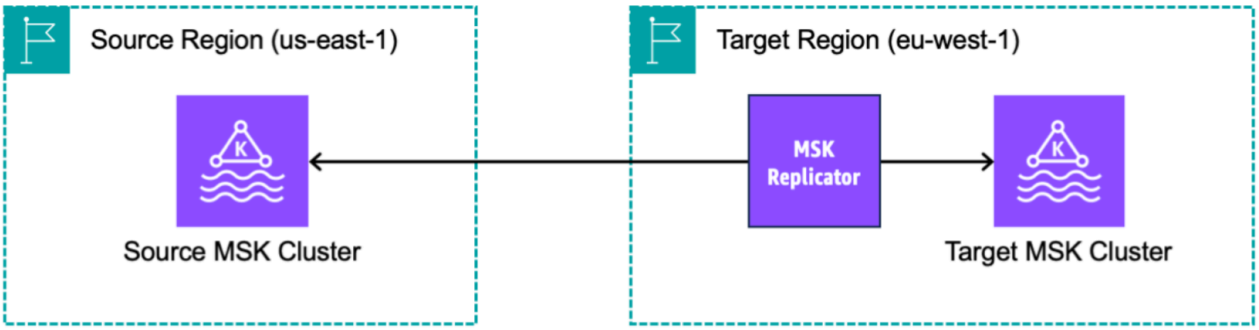
Il replicatore MSK non memorizza i dati. I dati vengono utilizzati dal cluster di origine, inseriti nel buffer in memoria e scritti nel cluster di destinazione. Il buffer viene cancellato automaticamente quando i dati vengono scritti correttamente o hanno esito negativo dopo nuovi tentativi. Tutte le comunicazioni e i dati tra MSK Replicator e i cluster sono sempre crittografati durante il transito. Tutte le chiamate API MSK Replicator, ad esempio`DescribeClusterV2`, vengono acquisite in. `CreateTopic` `DescribeTopicDynamicConfiguration` AWS CloudTrail Anche i log del vostro broker MSK rifletteranno la stessa cosa.

MSK Replicator crea argomenti nel cluster di destinazione con un fattore di replica pari a 3. Se necessario, è possibile modificare il fattore di replica direttamente sul cluster di destinazione.

Replica dei metadati

MSK Replicator supporta anche la copia dei metadati dal cluster di origine al cluster di destinazione. I metadati includono la configurazione degli argomenti, le liste di controllo degli accessi (ACL) di lettura e gli offset dei gruppi di consumatori. Come la replica dei dati, anche la replica dei metadati avviene in modo asincrono. Per prestazioni migliori, MSK Replicator dà priorità alla replica dei dati rispetto alla replica dei metadati.

Nell'ambito della sincronizzazione degli offset dei gruppi di consumatori, MSK Replicator effettua l'ottimizzazione per i consumatori del cluster di origine che leggono da una posizione più vicina alla fine del flusso (partizione di fine argomento). Se i gruppi di consumatori sono in ritardo rispetto al cluster di origine, è possibile riscontrare un ritardo maggiore per tali gruppi di consumatori sul cluster di destinazione rispetto a quello di origine. Ciò significa che, dopo il failover sul cluster di destinazione, i consumatori rielaboreranno più messaggi duplicati. Per ridurre questo ritardo, i tuoi utenti del cluster di origine dovrebbero recuperare il ritardo e iniziare a consumare dall'estremità dello stream (fine della partizione dell'argomento). Man mano che i consumatori recuperano il ritardo, MSK Replicator ridurrà automaticamente il ritardo.



Requisiti e considerazioni sulla creazione di un replicatore Amazon MSK

Prendi nota di questi requisiti del cluster MSK per l'esecuzione del replicatore Amazon MSK.

Argomenti

- Autorizzazioni necessarie per creare un replicatore MSK
- Tipi e versioni di cluster supportati
- Configurazione del cluster MSK Serverless
- Modifiche alla configurazione del cluster

Autorizzazioni necessarie per creare un replicatore MSK

Ecco un esempio della policy IAM necessaria per creare un replicatore MSK. L'operazione `kafka:TagResource` è necessaria solo durante la creazione del replicatore MSK vengono forniti dei tag.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "VisualEditor1",
            "Effect": "Allow",
```

```
            "Action": [
                "iam:PassRole",
                "iam:CreateServiceLinkedRole",
                "ec2:DescribeSubnets",
                "ec2:DescribeSecurityGroups",
                "ec2:CreateNetworkInterface",
                "ec2:DescribeVpcs",
                "kafka:CreateReplicator",
                "kafka:TagResource"
            ],
            "Resource": "*"
        }
    ]
}
```

Di seguito è riportata una policy IAM di esempio per descrivere il replicatore. È necessario specificare l'operazione `kafka:DescribeReplicator` o l'operazione `kafka:ListTagsForResource`, ma non entrambe.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "VisualEditor1",
            "Effect": "Allow",
            "Action": [
                "kafka:DescribeReplicator",
                "kafka:ListTagsForResource"
            ],
            "Resource": "*"
        }
    ]
}
```

Tipi e versioni di cluster supportati

Questi sono i requisiti per i tipi di istanze supportati, le versioni di Kafka e le configurazioni di rete.

- Il replicatore MSK supporta sia i cluster assegnati da MSK sia i cluster MSK Serverless in qualsiasi combinazione di cluster di origine e di destinazione. Al momento, il replicatore MSK non supporta altri tipi di cluster Kafka.

- I cluster MSK Serverless richiedono il Controllo degli accessi IAM, non supportano la replica delle ACL di Apache Kafka e offrono un supporto limitato per la replica della configurazione sull'argomento. Per informazioni, consulta MSK Serverless.
- MSK Replicator è supportato solo su cluster che eseguono Apache Kafka 2.7.0 o versioni successive, indipendentemente dal fatto che i cluster di origine e di destinazione si trovino nella stessa regione o in regioni diverse. AWS
- Il replicatore MSK supporta i cluster che utilizzano tipi di istanze m5.large o superiori. I cluster t3.small non sono supportati.
- Se si utilizza il replicatore MSK con un cluster assegnato da MSK, sono necessari almeno tre broker sia nel cluster di origine sia in quello di destinazione. È possibile replicare i dati tra cluster in due zone di disponibilità, ma in tali cluster sono necessari almeno quattro broker.
- I cluster MSK di origine e di destinazione devono appartenere allo stesso account. AWS La replica tra cluster in account diversi non è supportata.
- Se i cluster MSK di origine e di destinazione si trovano in AWS regioni diverse (interregionali), MSK Replicator richiede che il cluster di origine abbia la connettività privata multi-VPC attivata per il metodo IAM Access Control. Il multi-VPC non è richiesto per altri metodi di autenticazione sul cluster di origine. Il multi-VPC non è necessario se si replicano dati tra cluster nella stessa regione. AWS Per informazioni, consulta the section called “Connettività privata multi-VPC in un'unica regione”.

Configurazione del cluster MSK Serverless

- MSK Serverless supporta la replica di queste configurazioni di argomenti per i cluster di destinazione MSK Serverless durante la creazione degli argomenti: `cleanup.policy`, `compression.type`, `max.message.bytes`, `retention.bytes`, `retention.ms`.
- MSK Serverless supporta solo queste configurazioni degli argomenti durante la sincronizzazione della configurazione degli argomenti: `compression.type`, `max.message.bytes`, `retention.bytes`, `retention.ms`.
- Il replicatore utilizza 83 partizioni compattate sui cluster MSK Serverless di destinazione. Assicurati che i cluster MSK Serverless di destinazione abbiano un numero sufficiente di partizioni compattate. Per informazioni, consulta Quota di MSK Serverless.

Modifiche alla configurazione del cluster

- Si consiglia di non attivare o disattivare l'archiviazione a più livelli dopo la creazione del replicatore MSK. Se il cluster di destinazione non è a più livelli, MSK non copierà le configurazioni di archiviazione a più livelli, indipendentemente dal fatto che il cluster di origine sia a più livelli o meno. Se si attiva l'archiviazione a più livelli sul cluster di destinazione dopo la creazione del replicatore, è necessario ricreare il replicatore. Se si desidera copiare i dati da un cluster non a più livelli a un cluster a più livelli, non è necessario copiare le configurazioni degli argomenti. Consulta la sezione Abilitazione e disabilitazione dell'archiviazione a più livelli su un argomento esistente.
- Non modificare le impostazioni di configurazione del cluster dopo la creazione del replicatore MSK. Le impostazioni di configurazione del cluster vengono convalidate durante la creazione del replicatore MSK. Per evitare problemi con il replicatore MSK, non modificare le seguenti impostazioni dopo la creazione dello stesso.
  - Modifica il cluster MSK nel tipo di istanza t3.
  - Modifica i permessi del ruolo di esecuzione del servizio.
  - Disabilita la connettività privata multi-VPC di MSK.
  - Modifica la policy basata sulle risorse del cluster collegata.
  - Modifica le regole del gruppo di sicurezza del cluster.

Guida introduttiva all'utilizzo del replicatore Amazon MSK

Questo tutorial mostra come configurare un cluster di origine e un claster di destinazione nella stessa AWS regione o in regioni diverse. AWS Successivamente, puoi utilizzare questi cluster per creare un replicatore Amazon MSK.

Passaggio 1: preparazione del cluster di origine Amazon MSK

Se disponi già di un cluster di origine MSK creato per il replicatore MSK, assicurati che soddisfi i requisiti descritti in questa sezione. Altrimenti, segui questi passaggi per creare un cluster di origine serverless o assegnato da MSK.

Il processo per la creazione di un cluster di origine del replicatore MSK tra regioni e nella stessa regione è simile. Le differenze vengono evidenziate nelle seguenti procedure.

1. Crea un cluster MSK Serverless o assegnato con il Controllo degli accessi IAM attivato nella regione di origine. Il cluster di origine deve avere un minimo di tre broker.

2. Per un replicatore MSK tra regioni, se l'origine è un cluster assegnato, configuralo con la connettività privata multi-VPC attivata per gli schemi di Controllo degli accessi IAM. Tieni presente che il tipo di autenticazione non autenticato non è supportato quando è attivato il multi-VPC. Non è necessario attivare la connettività privata multi-VPC per altri schemi di autenticazione (mTLS o SASL/SCRAM). È possibile utilizzare contemporaneamente gli schemi di autenticazione mTLS o SASL/SCRAM per gli altri client che si connettono al cluster MSK. È possibile configurare la connettività privata multi-VPC tramite le Impostazioni di rete nei dettagli del cluster della console oppure tramite l'API `UpdateConnectivity`. Consulta la sezione Il proprietario del cluster attiva il multi-VPC. Se il cluster di origine è un cluster MSK Serverlesss, non è necessario attivare la connettività privata multi-VPC.

   Per un replicatore MSK nella stessa regione, il cluster di origine MSK non richiede una connettività privata multi-VPC e altri client possono comunque accedere al cluster utilizzando il tipo di autenticazione non autenticato.

3. Per i replicatori MSK tra regioni, è necessario collegare una policy di autorizzazione basata sulle risorse al cluster di origine. Ciò consente a MSK di connettersi a questo cluster per replicare i dati. È possibile eseguire questa operazione utilizzando le procedure CLI o AWS Console riportate di seguito. Consulta anche la sezione Policy basate sulle risorse di Amazon MSK. Non è necessario eseguire questa operazione per i replicatori MSK nella stessa regione.

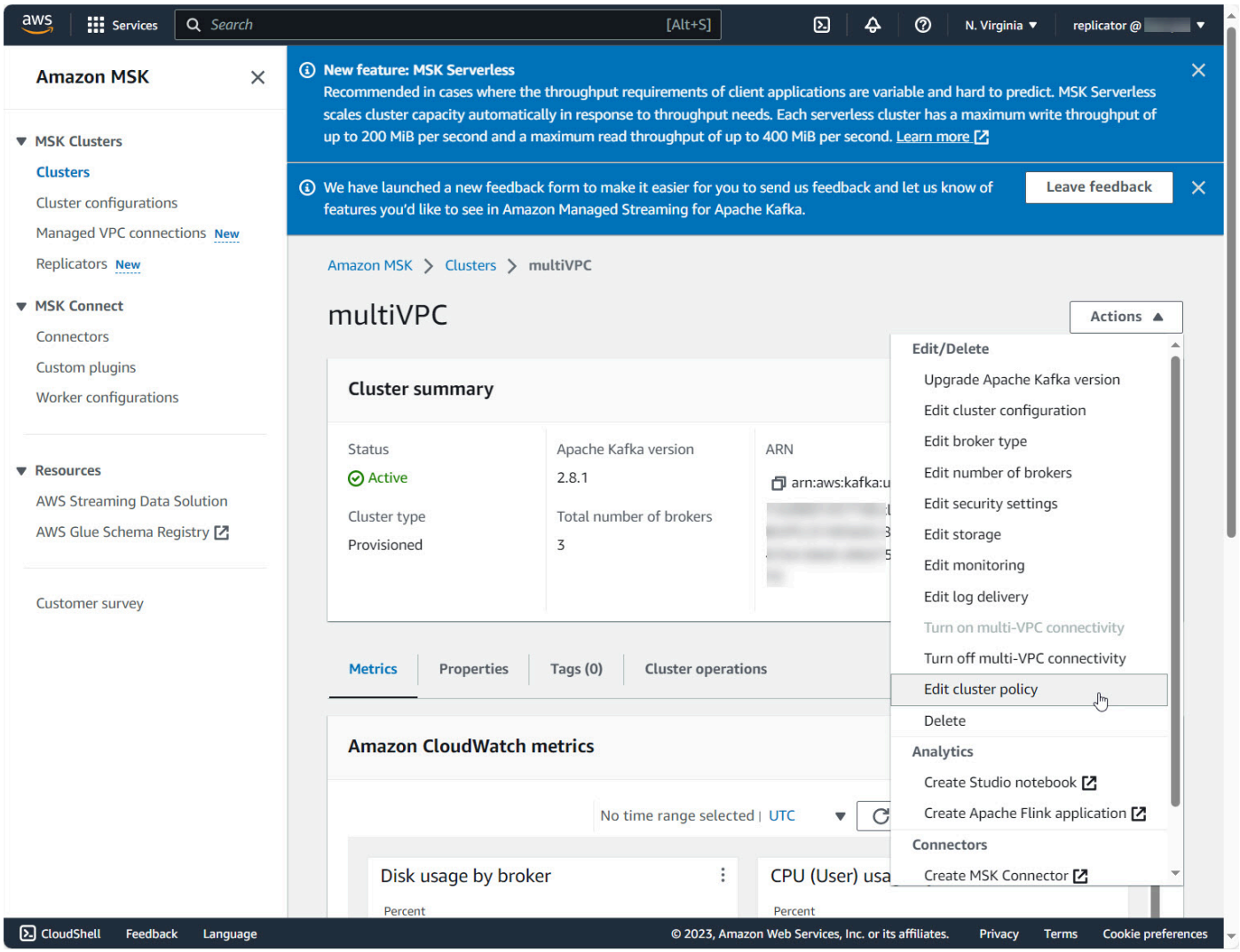
Console: create resource policy

Aggiorna la policy del cluster di origine con il seguente codice JSON. Sostituisci il segnaposto con l'ARN del tuo cluster di origine.

```
{
    "Version": "2012-10-17",
    "Statement": [
    {
        "Effect": "Allow",
        "Principal": {
            "Service": [
                "kafka.amazonaws.com"
            ]
        },
        "Action": [
            "kafka:CreateVpcConnection",
            "kafka:GetBootstrapBrokers",
            "kafka:DescribeClusterV2"
```

```
        ],
        "Resource": "<sourceClusterARN>"
    }
  ]
}
```

Utilizza l'opzione Modifica policy del cluster nel menu Operazioni nella pagina dei dettagli del cluster.



CLI: create resource policy

Nota: se utilizzi la AWS console per creare un cluster di origine e scegli l'opzione per creare un nuovo ruolo IAM, AWS allega la policy di fiducia richiesta al ruolo. Se invece desideri che MSK utilizzi un ruolo IAM esistente o se crei un ruolo autonomamente, collega la seguente policy di attendibilità a tale ruolo in modo che il replicatore MSK possa assumerlo. Per informazioni su come modificare la relazione di trust di un ruolo, consulta Modifica di un ruolo.

1. Recupera la versione corrente della policy del cluster MSK utilizzando questo comando. Sostituisci i segnaposto con l'ARN effettivo del cluster.

```
aws kafka get-cluster-policy –cluster-arn <Cluster ARN>
{
"CurrentVersion": "K1PA6795UKM GR7",
"Policy": "..."
}
```

2. Crea una policy basata sulle risorse per consentire al replicatore MSK di accedere al cluster di origine. Utilizza la seguente sintassi come modello sostituendo il segnaposto con l'ARN effettivo del cluster di origine.

```
aws kafka put-cluster-policy --cluster-arn "<sourceClusterARN>" --policy '{
"Version": "2012-10-17",
"Statement": [
{
"Effect": "Allow",
"Principal": {
"Service": [
"kafka.amazonaws.com"
]
},
"Action": [
"kafka:CreateVpcConnection",
"kafka:GetBootstrapBrokers",
"kafka:DescribeClusterV2"
],
"Resource": "<sourceClusterARN>"
}
]
```

Passaggio 2: preparazione del cluster di destinazione Amazon MSK

Crea un cluster di destinazione MSK (assegnato o serverless) con il Controllo degli accessi IAM attivato. Il cluster di destinazione non richiede l'attivazione della connettività privata multi-VPC. Il cluster di destinazione può trovarsi nella stessa AWS regione o in una regione diversa del cluster di origine. Sia il cluster di origine che quello di destinazione devono trovarsi nello stesso AWS account. Il cluster di destinazione deve avere un minimo di tre broker.

Passaggio 3: creazione di un replicatore Amazon MSK

Prima di creare un replicatore Amazon MSK, assicurati di disporre delle Autorizzazioni necessarie per creare un replicatore MSK.

Argomenti

- Creazione di un replicatore tramite la console AWS nella regione del cluster di destinazione
- Scelta del cluster di origine
- Scelta del cluster di destinazione
- Configurazione delle impostazioni e delle autorizzazioni del replicatore

Creazione di un replicatore tramite la console AWS nella regione del cluster di destinazione

1. Nella AWS regione in cui si trova il cluster MSK di destinazione, apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Scegli Replicatori per visualizzare l'elenco dei replicatori presenti nell'account.
3. Scegli Crea replicatore.
4. Nel riquadro Dettagli del replicatore, assegna un nome univoco al nuovo replicatore.

Scelta del cluster di origine

Il cluster di origine contiene i dati da copiare in un cluster MSK di destinazione.

1. Nel riquadro Cluster di origine, scegli la regione AWS in cui si trova il cluster di origine.

   È possibile cercare la regione di un cluster accedendo a Cluster MSK e controllando i dettagli dell'ARN in Cluster. Il nome della regione è incorporato nella stringa ARN. Nell'ARN di esempio seguente, il cluster primario si trova nella regione `ap-southeast-2`.

   ```
   arn:aws:kafka:ap-southeast-2:123456789012:cluster/cluster-11/
   eec93c7f-4e8b-4baf-89fb-95de01ee639c-s1
   ```

2. Inserisci l'ARN del tuo cluster di origine o naviga per scegliere il tuo cluster di origine.
3. Scegli una o più sottoreti per il cluster di origine.

La console mostra le sottoreti disponibili nella regione del cluster di origine che puoi selezionare. È necessario selezionare almeno due sottoreti. Per un replicatore MSK nella stessa regione, le sottoreti selezionate sono impostate per accedere al cluster di origine e le sottoreti per accedere al cluster di destinazione devono trovarsi nella stessa zona di disponibilità.

4. Scegli uno o più gruppi di sicurezza per il replicatore MSK per accedere al cluster di origine.

   - Per la replica tra regioni (CRR), non è necessario fornire gruppi di sicurezza per il cluster di origine.
   - Per la replica nella stessa regione (SRR), accedi alla console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/ e assicurati che i gruppi di sicurezza che fornirai per il Replicator dispongano di regole in uscita per consentire il traffico verso i gruppi di sicurezza del cluster di origine. Inoltre, assicurati che i gruppi di sicurezza del cluster di origine dispongano di regole in entrata che consentano il traffico proveniente dai gruppi di sicurezza Replicator forniti per l'origine.

     Per aggiungere regole in entrata al gruppo di sicurezza del cluster di origine:

     1. Nella AWS console, accedi ai dettagli del cluster di origine selezionando il nome del cluster.
     2. Seleziona la scheda Proprietà, quindi scorri verso il basso fino al riquadro Impostazioni di rete e seleziona il nome del Gruppo di sicurezza applicato.
     3. Vai alle regole in entrata e seleziona Modifica le regole in entrata.
     4. Seleziona Aggiungi regola.
     5. Nella colonna Tipo per la nuova regola, seleziona TCP personalizzato.
     6. Nella colonna Intervallo di porte, digita`9098`. MSK Replicator utilizza il controllo degli accessi IAM per connettersi al cluster che utilizza la porta 9098.
     7. Nella colonna Origine, digita il nome del gruppo di sicurezza che fornirai durante la creazione di Replicator per il cluster di origine (potrebbe essere lo stesso del gruppo di sicurezza del cluster di origine MSK), quindi seleziona Salva regole.

     Per aggiungere regole in uscita al gruppo di sicurezza di Replicator fornito per l'origine:

     1. Nella AWS console per Amazon EC2, vai al gruppo di sicurezza che fornirai durante la creazione di Replicator per l'origine.

2. Vai alle regole in uscita e seleziona Modifica regole in uscita.
3. Seleziona Aggiungi regola.
4. Nella colonna Tipo per la nuova regola, seleziona TCP personalizzato.
5. Nella colonna Intervallo di porte, digita`9098`. MSK Replicator utilizza il controllo degli accessi IAM per connettersi al cluster che utilizza la porta 9098.
6. Nella colonna Origine, digita il nome del gruppo di sicurezza del cluster di origine MSK, quindi seleziona Salva regole.

> **Note**
>
> In alternativa, se non desideri limitare il traffico utilizzando i tuoi gruppi di sicurezza, puoi aggiungere regole in entrata e in uscita che consentono All Traffic.
> 1. Seleziona Aggiungi regola.
> 2. Nella colonna Tipo, seleziona Tutto il traffico.
> 3. Nella colonna Origine, digita `0.0.0.0/0` e quindi seleziona Salva regole.

Scelta del cluster di destinazione

Il cluster di destinazione è il cluster MSK assegnato o serverless in cui vengono copiati i dati di origine.

> **Note**
>
> Il replicatore MSK crea nuovi argomenti nel cluster di destinazione con un prefisso generato automaticamente aggiunto al nome dell'argomento. Ad esempio, il replicatore MSK replica i dati in "`topic`" dal cluster di origine a un nuovo argomento nel cluster di destinazione denominato `<sourceKafkaClusterAlias>.topic`. Questo serve a distinguere gli argomenti che contengono i dati replicati dal cluster di origine da altri argomenti del cluster di destinazione ed evitare che i dati vengano replicati circolarmente tra i cluster. È possibile trovare il prefisso che verrà aggiunto ai nomi degli argomenti nel cluster di destinazione nel campo sourceKafkaClusterAlias utilizzando l'`DescribeReplicator`API o nella pagina dei dettagli del Replicator sulla console MSK. Il prefisso nel cluster di destinazione è < Alias>. sourceKafkaCluster

1. Nel riquadro Cluster di destinazione, scegli la AWS regione in cui si trova il cluster di destinazione.
2. Inserisci l'ARN del cluster di destinazione o sfoglia l'elenco per scegliere il cluster di destinazione.
3. Scegli una o più sottoreti per il tuo cluster di destinazione.

   La console mostra le sottoreti disponibili nella regione del cluster di destinazione che puoi selezionare. È necessario selezionare almeno due sottoreti.
4. Scegli uno o più gruppi di sicurezza per il replicatore MSK per accedere al cluster di destinazione.

   Vengono visualizzati i gruppi di sicurezza disponibili nella regione del cluster di destinazione che puoi selezionare. Il gruppo di sicurezza scelto è associato a ciascuna connessione. Per ulteriori informazioni sull'uso dei gruppi di sicurezza, consulta la sezione Controlla il traffico verso AWS le tue risorse utilizzando i gruppi di sicurezza nella Amazon VPC User Guide.

   - Sia per la replica tra regioni (CRR) che per la replica su stessa regione (SRR), accedi alla console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/ e assicurati che i gruppi di sicurezza che fornirai al Replicator dispongano di regole in uscita per consentire il traffico verso i gruppi di sicurezza del cluster di destinazione. Inoltre, assicurati che i gruppi di sicurezza del cluster di destinazione dispongano di regole in entrata che consentano il traffico verso i gruppi di sicurezza del replicatore forniti per la destinazione.

     Per aggiungere regole in entrata al gruppo di sicurezza del cluster di destinazione:

     1. Nella AWS console, accedi ai dettagli del cluster di destinazione selezionando il nome del cluster.
     2. Seleziona la scheda Proprietà, quindi scorri verso il basso fino al riquadro Impostazioni di rete per selezionare il nome del gruppo di sicurezza applicato.
     3. Vai alle regole in entrata e seleziona Modifica le regole in entrata.
     4. Seleziona Aggiungi regola.
     5. Nella colonna Tipo per la nuova regola, seleziona TCP personalizzato.
     6. Nella colonna Intervallo di porte, digita`9098`. MSK Replicator utilizza il controllo degli accessi IAM per connettersi al cluster che utilizza la porta 9098.

7. Nella colonna Origine, digita il nome del gruppo di sicurezza che fornirai durante la creazione di Replicator per il cluster di destinazione (potrebbe essere lo stesso del gruppo di sicurezza del cluster di destinazione MSK), quindi seleziona Salva regole.

Per aggiungere regole in uscita al gruppo di sicurezza di Replicator fornito per la destinazione:

1. Nella AWS console, vai al gruppo di sicurezza che fornirai durante la creazione di Replicator per la destinazione.
2. Seleziona la scheda Proprietà, quindi scorri verso il basso fino al riquadro Impostazioni di rete per selezionare il nome del gruppo di sicurezza applicato.
3. Vai alle regole in uscita e seleziona Modifica regole in uscita.
4. Seleziona Aggiungi regola.
5. Nella colonna Tipo per la nuova regola, seleziona TCP personalizzato.
6. Nella colonna Intervallo di porte, digita`9098`. MSK Replicator utilizza il controllo degli accessi IAM per connettersi al cluster che utilizza la porta 9098.
7. Nella colonna Origine, digita il nome del gruppo di sicurezza del cluster di destinazione MSK, quindi seleziona Salva regole.

> **Note**
>
> In alternativa, se non desideri limitare il traffico utilizzando i tuoi gruppi di sicurezza, puoi aggiungere regole in entrata e in uscita che consentono All Traffic.
> 1. Seleziona Aggiungi regola.
> 2. Nella colonna Tipo, seleziona Tutto il traffico.
> 3. Nella colonna Origine, digita `0.0.0.0/0` e quindi seleziona Salva regole.

Configurazione delle impostazioni e delle autorizzazioni del replicatore

1. Nel riquadro Impostazioni del replicatore, specifica gli argomenti che desideri replicare utilizzando le espressioni regolari negli elenchi consentiti e non consentiti. Come impostazione predefinita, vengono replicati tutti gli argomenti.

> **Note**
>
> MSK Replicator replica solo fino a 750 argomenti in ordine ordinato. Se è necessario replicare più argomenti, si consiglia di creare un Replicator separato. Vai al Support Center della AWS console e crea un caso di supporto se hai bisogno di supporto per più di 750 argomenti per Replicator. È possibile monitorare il numero di argomenti replicati utilizzando la metrica TopicCount "». Per informazioni, consulta Quota di Amazon MSK.

2. Per impostazione predefinita, MSK Replicator avvia la replica dall'offset più recente (più recente) negli argomenti selezionati. In alternativa, è possibile avviare la replica dal primo offset (il più vecchio) negli argomenti selezionati se si desidera replicare i dati esistenti sugli argomenti. Una volta creato il replicatore, non è possibile modificare questa impostazione. Questa impostazione corrisponde al `startingPosition`campo delle API di `CreateReplicator`richiesta e `DescribeReplicator`risposta.

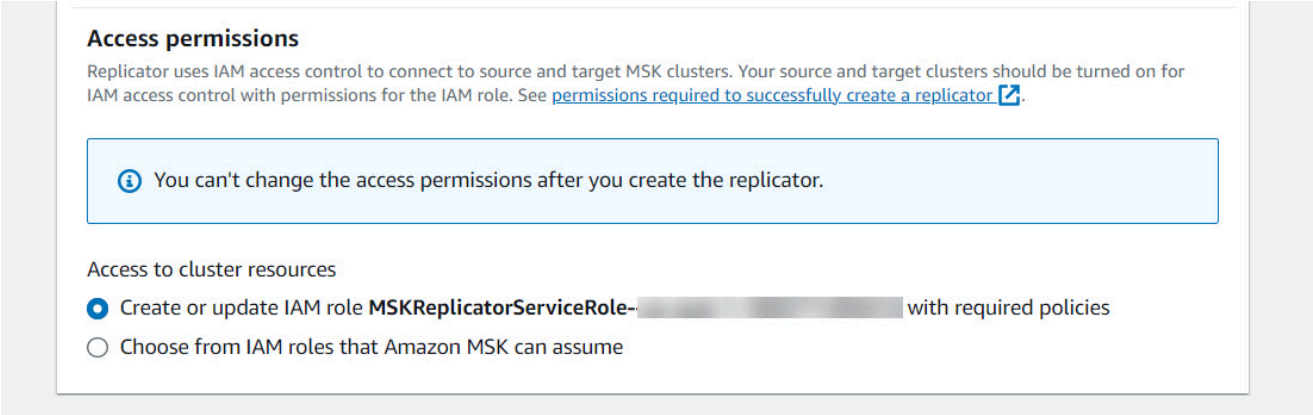
   > **Note**
   >
   > MSK Replicator agisce come un nuovo consumatore per il cluster di origine. A seconda della quantità di dati che si stanno replicando e della capacità di consumo disponibile nel cluster di origine, ciò può causare limitazioni agli altri consumatori del cluster di origine. Se si crea un Replicator impostato sulla prima posizione iniziale, MSK Replicator leggerà una serie di dati all'inizio, che potrebbe consumare tutta la capacità di consumo del cluster di origine. Una volta che il Replicator avrà recuperato il ritardo, il tasso di consumo dovrebbe diminuire per corrispondere al throughput degli argomenti del cluster di origine. Se si esegue la replica dalla prima posizione, si consiglia di gestire il throughput di Replicator utilizzando le quote Kafka per evitare che altri consumatori subiscano limitazioni.

3. Per impostazione predefinita, il replicatore MSK copia tutti i metadati, incluse le configurazioni degli argomenti, le liste di controllo degli accessi (ACL) e gli offset dei gruppi di consumatori per un failover senza interruzioni. Se non si sta creando il replicatore per il failover, è possibile scegliere facoltativamente di disattivare una o più di queste impostazioni disponibili nella sezione Impostazioni aggiuntive.

> **Note**
>
> Il replicatore MSK non replica le ACL di scrittura poiché i produttori non dovrebbero scrivere direttamente sull'argomento replicato nel cluster di destinazione. I produttori devono scrivere sull'argomento locale nel cluster di destinazione dopo il failover. Per informazioni dettagliate, vedi Esecuzione di un failover pianificato nella regione secondaria AWS.

4. Nel riquadro Replica del gruppo di consumatori, specifica i gruppi di consumatori che desideri replicare utilizzando le espressioni regolari negli elenchi consentiti e non consentiti. Per impostazione predefinita, vengono replicati tutti i gruppi di consumatori.
5. Nel riquadro Compressione, puoi facoltativamente scegliere di comprimere i dati scritti nel cluster di destinazione. Se intendi utilizzare la compressione, ti consigliamo di utilizzare lo stesso metodo di compressione dei dati nel cluster di origine.
6. Nel riquadro Autorizzazioni di accesso, effettua una delle operazioni seguenti:

   a. Seleziona Crea o aggiorna il ruolo IAM con le policy richieste. La console MSK collegherà automaticamente le autorizzazioni e la policy di attendibilità necessarie al ruolo di esecuzione del servizio richiesto per la lettura e la scrittura nei cluster MSK di origine e di destinazione.



   b. Fornisci il tuo ruolo IAM selezionando Scegli tra i ruoli IAM che Amazon MSK può assumere. Ti consigliamo di collegare la policy IAM `AWSMSKReplicatorExecutionRole` gestita al tuo ruolo di esecuzione del servizio, anziché scrivere la tua politica IAM.

      - Crea il ruolo IAM che il replicatore utilizzerà per leggere e scrivere nei cluster MSK di origine e di destinazione utilizzando il codice JSON riportato di seguito come parte della policy di attendibilità e la `AWSMSKReplicatorExecutionRole` collegata al

ruolo. Nella policy di attendibilità, sostituisci il segnaposto <yourAccountID> con l'ID dell'account effettivo.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "kafka.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "<yourAccountID>"
                }
            }
        }
    ]
}
```

7. Nel riquadro Tag del replicatore, puoi facoltativamente assegnare tag alla risorsa replicatore MSK. Per ulteriori informazioni, consulta Assegnazione di tag a un cluster Amazon MSK. Per un replicatore MSK tra regioni, i tag vengono sincronizzati automaticamente con la regione remota al momento della creazione del replicatore. Se si modificano i tag dopo la creazione del replicatore, la modifica non viene sincronizzata automaticamente con la regione remota, quindi sarà necessario sincronizzare manualmente i riferimenti del replicatore locale e del replicatore remoto.
8. Seleziona Crea.

Se desideri limitare le `kafka-cluster:WriteData` autorizzazioni, consulta la sezione Creare politiche di autorizzazione di Come funziona il controllo degli accessi IAM per Amazon MSK. Dovrai aggiungere l'`kafka-cluster:WriteDataIdempotently`autorizzazione sia al cluster di origine che a quello di destinazione.

La creazione e la transizione del replicatore MSK allo stato RUNNING richiedono circa 30 minuti.

Se si crea un nuovo replicatore MSK per sostituirne uno eliminato, il nuovo replicatore avvia la replica dall'offset più recente.

Se il replicatore MSK è passato allo stato FAILED, consulta la sezione Risoluzione dei problemi relativi al replicatore MSK.

Modifica delle impostazioni del replicatore MSK

Non è possibile modificare il cluster di origine, il cluster di destinazione o la posizione iniziale del replicatore una volta creato MSK Replicator. Tuttavia, è possibile modificare altre impostazioni di Replicator, ad esempio argomenti e gruppi di consumatori, da replicare.

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Nel riquadro di navigazione a sinistra, scegli Replicatori per visualizzare l'elenco dei replicatori presenti nell'account e seleziona il replicatore MSK che desideri modificare.
3. Scegliere la scheda Properties (Proprietà).
4. Nella sezione Impostazioni del replicatore, scegli Modifica replicatore.
5. È possibile modificare le impostazioni del replicatore MSK modificando una qualsiasi di queste impostazioni.

   - Specifica gli argomenti che desideri replicare utilizzando le espressioni regolari negli elenchi consentiti e non consentiti. Per impostazione predefinita, il replicatore MSK copia tutti i metadati, incluse le configurazioni degli argomenti, le liste di controllo degli accessi (ACL) e gli offset dei gruppi di consumatori per un failover senza interruzioni. Se non si sta creando il replicatore per il failover, è possibile scegliere facoltativamente di disattivare una o più di queste impostazioni disponibili nella sezione Impostazioni aggiuntive.

     > **Note**
     >
     > Il replicatore MSK non replica le ACL di scrittura poiché i produttori non dovrebbero scrivere direttamente sull'argomento replicato nel cluster di destinazione. I produttori devono scrivere sull'argomento locale nel cluster di destinazione dopo il failover. Per informazioni dettagliate, vedi Esecuzione di un failover pianificato nella regione secondaria AWS.

   - In Replica del gruppo di consumatori, puoi specificare i gruppi di consumatori che desideri replicare utilizzando le espressioni regolari negli elenchi consentiti e non consentiti. Per

impostazione predefinita, vengono replicati tutti i gruppi di consumatori. Se gli elenchi consentiti e non consentiti sono vuoti, la replica dei gruppi di consumatori è disattivata.

- In Tipo di compressione di destinazione, puoi scegliere se comprimere i dati scritti nel cluster di destinazione. Se intendi utilizzare la compressione, ti consigliamo di utilizzare lo stesso metodo di compressione dei dati nel cluster di origine.

6. Salvare le modifiche.

   La creazione e la transizione del replicatore MSK allo stato di esecuzione richiedono circa 30 minuti. Se il replicatore MSK è passato allo stato FAILED, consulta la sezione ??? sulla risoluzione dei problemi.

Eliminazione di un replicatore MSK

Se la creazione di un replicatore MSK ha esito negativo (stato FAILED), potrebbe essere necessario eliminarlo. I cluster di origine e di destinazione assegnati a un replicatore MSK non possono essere modificati una volta creato il replicatore MSK. È possibile eliminare un replicatore MSK esistente e crearne uno nuovo. Se si crea un nuovo replicatore MSK per sostituire quello eliminato, il nuovo replicatore avvia la replica dall'offset più recente.

1. Nella AWS regione in cui si trova il cluster di origine, accedi e apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/. AWS Management Console
2. Nel riquadro di navigazione, seleziona Replicatori.
3. Dall'elenco dei replicatori MSK, seleziona quello che desideri eliminare e scegli Elimina.

Monitoraggio della replica

È possibile utilizzare https://console.aws.amazon.com/cloudwatch/ nella regione del cluster di destinazione per visualizzare i parametri per `ReplicationLatency`, `MessageLag` e `ReplicatorThroughput` a livello di argomento e aggregati per ogni replicatore Amazon MSK. Le metriche sono visibili sotto lo spazio dei ReplicatorNamenomi «AWS/Kafka». Per verificare la presenza di problemi, puoi anche consultare i parametri `ReplicatorFailure`, `AuthError` e `ThrottleTime`.

La console MSK visualizza un sottoinsieme di metriche per ogni MSK Replicator. CloudWatch Dall'elenco dei Replicatori della console, seleziona il nome di un replicatore e scegli la scheda Monitoraggio.

Parametri del replicatore MSK

I parametri seguenti descrivono i parametri delle prestazioni o delle connessioni per il replicatore MSK.

AuthError le metriche non coprono gli errori di autenticazione a livello di argomento. Per monitorare gli errori di autenticazione a livello di argomento di MSK Replicator, monitorate le metriche di Replicator e le metriche a livello di argomento del ReplicationLatency cluster di origine,. MessagesInPerSec Se un argomento è ReplicationLatency stato ridotto a 0 ma l'argomento contiene ancora dati in corso, significa che il Replicator ha un problema di autenticazione con l'argomento. Verifica che il ruolo IAM per l'esecuzione del servizio del replicatore disponga di autorizzazioni sufficienti per accedere all'argomento.

Tipo di parametr	Parametro	Descrizione	Dimensio i	Unità	Granular tà dei parametr grezzi	Statistic he di aggregaz one dei parametr grezzi	
Prestazio ni	Replicati onLatency	Tempo impiegato dai record per la replica dal cluster di origine a quello di destinazione; tempo che intercorre tra l'ora di produzion e di un record all'origine e l'ora di replica alla	Replicato rName	Milliseco ndi	Partizion e	Massimo	
			Replicato rName, Argomer	Milliseco ndi	Partizion e	Massimo	

Tipo di parametr	Parametro	Descrizione	Dimensio i	Unità	Granular tà dei parametr grezzi	Statistic he di aggrega one dei parametr grezzi	
		destinazione. Se ReplicationLatency aumenta, controlla se i cluster hanno partizioni sufficienti per supportare la replica. Una latenza di replica elevata può verificarsi quando il numero di partizioni è troppo basso per una velocità di trasmissione effettiva elevata.					

Tipo di parametr	Parametro	Descrizione	Dimensi i	Unità	Granular tà dei parametr grezzi	Statistic he di aggrega one dei parametr grezzi
Prestazic ni	MessageLag	Monitora la sincroniz zazione tra MSK Replicator e il cluster di origine. MessageLag indica il ritardo tra i messaggi prodotti nel cluster di origine e i messaggi consumati dal replicatore. Non è il ritardo tra il cluster di origine e quello di destinazi one. Anche se il cluster di origine non è disponibi le/interrotto, il replicatore finirà di scrivere il messaggio che ha utilizzat o nel cluster di destinazi one. Dopo	Replicato rName	Conteggi	Partizion e	Somma
			Replicato rName, Argomer	Conteggi	Partizion e	Somma

Tipo di parametr	Parametro	Descrizione	Dimensi i	Unità	Granular tà dei parametr grezzi	Statistic he di aggrega one dei parametr grezzi	
		un'interruzione, MessageLa g mostra un aumento che indica il numero di messaggi che il replicatore si trova dietro al cluster di origine e questo può essere monitorat o fino a quando il numero di messaggi non raggiunge 0, a dimostrazione del fatto che il replicatore ha raggiunto il cluster di origine.					

Tipo di parametr	Parametro	Descrizione	Dimensio i	Unità	Granular tà dei parametr grezzi	Statistic he di aggrega one dei parametr grezzi	
Prestazio ni	Replicato rThroughput	Numero medio di byte replicati al secondo. Se si Replicato rThroughput tratta di un argomento, di un controllo KafkaClus terPingSu ccessCount e di AuthError parametri per garantire che il Replicator sia in grado di comunicare con i cluster, controlla te i parametri del cluster per assicurarvi che il cluster non sia inattivo.	Replicato rName	BytesPer econd	Partizion e	Somma	
			Replicato rName, Argomer	BytesPer econd	Partizion e	Somma	

Tipo di parametr	Parametro	Descrizione	Dimensi i	Unità	Granular tà dei parametr grezzi	Statistic he di aggrega one dei parametr grezzi
Esegui il debug	AuthError	Il numero di connessioni con autenticazione non riuscita al secondo. Se questo parametro è superiore a 0, puoi verificar e se la policy del ruolo di esecuzione del servizio per il replicato re è valida e assicurarti che non siano impostate autorizzazioni di rifiuto per le autorizza zioni del cluster. In base alla dimensione clusterAlias, è possibile verificar e se è il cluster di origine o di	Replicato rName, ClusterA ias	Conteggi	Worker	Somma

Tipo di parametr	Parametro	Descrizione	Dimensi i	Unità	Granular tà dei parametr grezzi	Statistic he di aggrega one dei parametr grezzi	
		destinazione a presentare errori di autentica zione.					

Tipo di parametr	Parametro	Descrizione	Dimensi i	Unità	Granular tà dei parametr grezzi	Statistic he di aggrega one dei parametr grezzi
Esegui il debug	ThrottleTime	Il tempo medio, espresso in millisecondi, per il quale i broker del cluster hanno limitato la larghezza di banda della rete per una richiesta . Imposta la limitazione della larghezza di banda della rete per evitare che il replicatore MSK sovraccar ichi il cluster. Se questo parametro è 0, replicati onLatency non è elevato e replicato rThroughp ut è come previsto, allora la limitazione della larghezza	Replicato rName, ClusterA ias	Milliseco ndi	Worker	Massimo

Tipo di parametr	Parametro	Descrizione	Dimensi i	Unità	Granular tà dei parametr grezzi	Statistic he di aggrega one dei parametr grezzi	
		di banda della rete funziona come previsto. Se questo parametro è superiore a 0, è possibile regolare la limitazione della larghezza di banda della rete di conseguenza.					
Esegui il debug	ReplicatorFailure	Numero di errori riscontrati dal replicatore.	Replicat rName	Conteggi		Somma	

Tipo di parametr	Parametro	Descrizione	Dimensi i	Unità	Granular tà dei parametr grezzi	Statistic he di aggregaz one dei parametr grezzi	
Esegui il debug	KafkaClus terPingSu ccessCount	Indica lo stato della connessio ne del replicatore al cluster Kafka. Se questo valore è 1, la connessio ne è integra. Se il valore è 0 o nessun punto di dati, la connessione non è integra. Se il valore è 0, puoi controlla re le impostazi oni di rete o di autorizzazione IAM per il cluster Kafka. In base alla ClusterAlias dimensione, è possibile identific are se questa metrica è per il cluster di origine o di destinazi one.	Replicato rName, ClusterA ias	Conteggi		Somma	

Utilizzo della replica per aumentare la resilienza di un'applicazione di streaming Kafka tra regioni

È possibile utilizzare MSK Replicator per configurare topologie di cluster attive-attive o attive-passive per aumentare la resilienza dell'applicazione Apache Kafka in tutte le regioni. AWS In una configurazione attiva-attiva, entrambi i cluster MSK eseguono attivamente operazioni di lettura e scrittura. In una configurazione attiva-passiva, solo un cluster MSK alla volta serve attivamente lo streaming di dati, mentre l'altro cluster è in standby.

Considerazioni sulla creazione di applicazioni Apache Kafka tra regioni

I consumatori devono essere in grado di rielaborare i messaggi duplicati senza ripercussioni a valle. MSK Replicator replica i dati che possono causare duplicati nel cluster di standby. at-least-once Quando si passa alla AWS regione secondaria, i consumatori possono elaborare gli stessi dati più di una volta. Il replicatore MSK dà la priorità alla copia dei dati rispetto agli offset dei consumatori per prestazioni migliori. Dopo un failover, il consumatore può iniziare a leggere dagli offset precedenti con conseguente duplicazione dell'elaborazione.

I produttori e i consumatori devono inoltre tollerare una perdita minima di dati. Poiché MSK Replicator replica i dati in modo asincrono, quando la AWS regione primaria inizia a riscontrare errori, non vi è alcuna garanzia che tutti i dati vengano replicati nella regione secondaria. È possibile utilizzare la latenza di replica per determinare il numero massimo di dati che non sono stati copiati nella regione secondaria.

Utilizzo della topologia di cluster attiva-attiva rispetto a quella attiva-passiva

Una topologia di cluster attiva-attiva offre tempi di ripristino quasi nulli e la possibilità per l'applicazione di streaming di funzionare contemporaneamente in più regioni AWS . Quando un cluster in una regione è danneggiato, le applicazioni connesse al cluster nell'altra regione continuano a elaborare i dati.

Le configurazioni attive-passive sono adatte alle applicazioni che possono essere eseguite in una sola regione AWS alla volta o quando è necessario un maggiore controllo sull'ordine di elaborazione dei dati. Le configurazioni attive-passive richiedono più tempo di ripristino rispetto alle configurazioni attive-attive, poiché, dopo un failover, è necessario avviare l'intera configurazione attiva-passiva, compresi produttori e consumatori, nella regione secondaria per riprendere lo streaming dei dati.

Creazione di una configurazione di cluster Kafka attiva-passiva e denominazione replicata degli argomenti

Per una configurazione attiva-passiva, si consiglia di utilizzare una configurazione simile per produttori, cluster MSK e consumatori (con lo stesso nome di gruppo di consumatori) in due regioni diverse. AWS È importante che i due cluster MSK abbiano capacità di lettura e scrittura identiche per garantire una replica affidabile dei dati. È necessario creare un replicatore MSK per copiare continuamente i dati dal cluster primario a quello di standby. È inoltre necessario configurare i produttori per scrivere i dati in argomenti su un cluster nella stessa regione. AWS

Per garantire che i consumatori possano riavviare in modo affidabile l'elaborazione dal cluster di standby, è necessario configurare i consumatori in modo che leggano i dati dagli argomenti utilizzando un carattere jolly ".*". Ad esempio, MSK Replicator replica «topic1" dal cluster primario in un nuovo argomento nel cluster di standby chiamato «< alias>.topic1». sourceKafkaCluster Ad esempio, è possibile configurare i produttori in modo che scrivano su "topic1" e i consumatori in modo che utilizzino ".*topic1" in entrambe le regioni. Questo esempio includerebbe anche un argomento come footopic1, quindi modifica l'operatore jolly in base alle tue esigenze.

AWS Quando eseguire il failover nella regione secondaria

Si consiglia di monitorare la latenza di replica nella regione secondaria AWS utilizzando. CloudWatch Durante un evento di servizio nella AWS regione principale, la latenza di replica può aumentare improvvisamente. Se la latenza continua ad aumentare, utilizza il AWS Service Health Dashboard per verificare la presenza di eventi di servizio nella AWS regione principale. Se si verifica un evento, puoi eseguire il failover nella regione secondaria AWS .

Esecuzione di un failover pianificato nella regione secondaria AWS

È possibile eseguire un failover pianificato per testare la resilienza dell'applicazione rispetto a un evento imprevisto nella AWS regione principale in cui si trova il cluster MSK di origine. Un failover pianificato non dovrebbe comportare la perdita di dati.

1. Arresta tutti i produttori e i consumatori che si connettono al cluster di origine.
2. Crea un nuovo replicatore MSK per replicare i dati dal cluster MSK nella regione secondaria al cluster MSK nella regione primaria. Ciò è necessario per copiare i dati che verranno scritti dalla regione secondaria alla regione primaria, in modo da poter eseguire il failback nella regione primaria al termine dell'evento imprevisto.

3. Avvia i produttori sul cluster bersaglio nella AWS regione secondaria.
4. A seconda dei requisiti di ordinamento dei messaggi dell'applicazione, segui i passaggi indicati in una delle schede seguenti.

No message ordering

Se l'applicazione non richiede l'ordinamento dei messaggi, avvia i consumatori della AWS regione secondaria che leggono sia gli argomenti locali (ad esempio`topic`) che quelli replicati (ad esempio`<sourceKafkaClusterAlias>.topic`) utilizzando un operatore wildcard (ad esempio,. *argomento).

Message ordering

Se l'applicazione richiede l'ordinamento dei messaggi, avvia i consumatori solo per gli argomenti replicati sul cluster di destinazione (ad esempio, `<sourceKafkaClusterAlias>.topic`) ma non per gli argomenti locali (ad esempio, `topic`).

1. Attendi che tutti i consumatori degli argomenti replicati sul cluster MSK di destinazione completino l'elaborazione di tutti i dati, in modo che il ritardo del consumatore sia 0 e anche il numero di record elaborati sia 0. Quindi, interrompi i consumatori per gli argomenti replicati sul cluster di destinazione. A questo punto, tutti i record replicati dal cluster MSK di origine al cluster MSK di destinazione sono stati utilizzati.
2. Avvia i consumatori per gli argomenti locali (ad esempio, `topic`) sul cluster MSK di destinazione.

Esecuzione di un failover non pianificato nella regione secondaria AWS

È possibile eseguire un failover non pianificato quando si verifica un evento di servizio nella AWS regione principale in cui si trova il cluster MSK di origine e si desidera reindirizzare temporaneamente il traffico verso la AWS regione secondaria che ha il cluster MSK di destinazione. Un failover non pianificato potrebbe causare una perdita di dati.

1. Prova a chiudere tutti i produttori e i consumatori che si connettono al cluster MSK di origine nella regione primaria. Questa operazione potrebbe avere esito negativo.
2. Avvia i produttori che si connettono al cluster MSK di destinazione nella regione secondaria.
3. A seconda dei requisiti di ordinamento dei messaggi dell'applicazione, segui i passaggi indicati in una delle schede seguenti.

No message ordering

Se l'applicazione non richiede l'ordinamento dei messaggi, avviate i consumatori della AWS regione di destinazione che leggono sia gli argomenti locali (ad esempio`topic`) che quelli replicati (ad esempio) utilizzando un operatore wildcard (ad esempio,`<sourceKafkaClusterAlias>.topic`). `.*topic`

Message ordering

1. Avvia i consumatori solo per gli argomenti replicati nel cluster di destinazione (ad esempio, `<sourceKafkaClusterAlias>.topic`) ma non per gli argomenti locali (ad esempio, `topic`).
2. Attendi che tutti i consumatori degli argomenti replicati sul cluster MSK di destinazione completino l'elaborazione di tutti i dati, in modo che il ritardo di offset sia 0 e anche il numero di record elaborati sia 0. Quindi, interrompi i consumatori per gli argomenti replicati sul cluster di destinazione. A questo punto, tutti i record replicati dal cluster MSK di origine al cluster MSK di destinazione sono stati utilizzati.
3. Avvia i consumatori per gli argomenti locali (ad esempio, `topic`) sul cluster MSK di destinazione.

4. Una volta terminato l'evento di servizio nella regione primaria, create un nuovo replicatore MSK per replicare i dati dal cluster MSK nella regione secondaria al cluster MSK nella regione primaria con la posizione iniziale del replicatore impostata sulla prima. Ciò è necessario per copiare i dati che verranno scritti dalla regione secondaria alla regione primaria, in modo da poter eseguire il failback nella regione primaria al termine dell'evento di servizio. Se non impostate la posizione iniziale del Replicator sulla prima, i dati prodotti nel cluster nella regione secondaria durante l'evento di servizio nell'area primaria non verranno copiati nuovamente nel cluster nell'area primaria.

Esecuzione del failback nella regione primaria AWS

È possibile eseguire il failback AWS nella regione principale al termine dell'evento di servizio in tale regione. Il replicatore MSK salta automaticamente gli argomenti che hanno l'alias del cluster di origine come prefisso mentre esegue la replica dei dati nella regione primaria durante il failback.

Se avete seguito i passaggi di failover non pianificati, dovreste aver già creato il failback Replicator come parte dell'ultimo passaggio del failover dalla regione primaria a quella secondaria.

Se non hai seguito i passaggi di failover non pianificati, una volta terminato l'evento di servizio nella regione primaria, crea un nuovo replicatore MSK per replicare i dati dal tuo cluster MSK nella regione secondaria al tuo cluster MSK nella regione primaria con la posizione iniziale del replicatore impostata sulla prima. Ciò è necessario per copiare i dati che verranno scritti dalla regione secondaria alla regione primaria, in modo da poter eseguire il failback nella regione primaria al termine dell'evento di servizio. Se non si modifica la posizione iniziale del Replicator dal valore predefinito di più recente a più recente, i dati prodotti nel cluster nella regione secondaria durante l'evento di servizio nell'area primaria non verranno copiati nuovamente nel cluster nell'area primaria.

È necessario avviare le fasi di failback solo dopo che la replica dal cluster nella regione secondaria al cluster nella regione primaria è stata completata e la MessageLag metrica in è vicina a 0. CloudWatch Un failback pianificato non dovrebbe comportare la perdita di dati.

1. Arresta tutti i produttori e i consumatori che si connettono al cluster MSK nella regione secondaria.
2. Per la topologia attiva-passiva, elimina il replicatore che replica i dati dal cluster nella regione secondaria alla regione primaria. Non è necessario eliminare il replicatore per la topologia attiva-attiva.
3. Avvia i produttori che si connettono al cluster MSK nella regione primaria.
4. A seconda dei requisiti di ordinamento dei messaggi dell'applicazione, segui i passaggi indicati in una delle schede seguenti.

   No message ordering

   Se l'applicazione non richiede l'ordinamento dei messaggi, avviate i consumatori della AWS regione primaria che leggono sia gli argomenti locali (ad esempio`topic`) che quelli replicati (ad esempio) utilizzando un operatore wildcard (ad esempio,`<sourceKafkaClusterAlias>.topic`). `.*topic` I consumatori sugli argomenti locali (ad esempio: topic) riprenderanno dall'ultimo offset utilizzato prima del failover. Se erano presenti dati non elaborati prima del failover, verranno elaborati ora. Nel caso di un failover pianificato, tale record non dovrebbe esistere.

   Message ordering

   1. Avvia i consumatori solo per gli argomenti replicati nella regione primaria (ad esempio, `<sourceKafkaClusterAlias>.topic`) ma non per gli argomenti locali (ad esempio, `topic`).

2. Attendi che tutti i consumatori degli argomenti replicati sul cluster nella regione primaria completino l'elaborazione di tutti i dati, in modo che il ritardo di offset sia 0 e anche il numero di record elaborati sia 0. Quindi, interrompi i consumatori per gli argomenti replicati sul cluster nella regione primaria. A questo punto, tutti i record prodotti nella regione secondaria dopo il failover sono stati utilizzati nella regione primaria.
3. Avvia i consumatori per gli argomenti locali (ad esempio, `topic`) sul cluster nella regione primaria.

5. Verificate che il replicatore esistente dal cluster nella regione primaria al cluster nella regione secondaria sia nello stato RUNNING e funzioni come previsto utilizzando le `ReplicatorThroughput` metriche di latenza e.

Creazione di una configurazione attiva-attiva utilizzando il replicatore MSK

Segui questi passaggi per configurare la topologia attiva-attiva tra il cluster MSK di origine A e il cluster MSK di destinazione B.

1. Crea un replicatore MSK con il cluster MSK A come origine e il cluster MSK B come destinazione.
2. Dopo aver creato correttamente il replicatore MSK precedente, crea un replicatore con il cluster B come origine e il cluster A come destinazione.
3. Crea due set di produttori, ognuno dei quali scrive i dati contemporaneamente nell'argomento locale (ad esempio, "topic") nel cluster nella stessa regione del produttore.
4. Crea due set di consumatori, ciascuno dei quali legge i dati utilizzando un abbonamento wildcard (ad esempio». *topic») dal cluster MSK nella stessa AWS regione del consumatore. In questo modo, i consumatori leggeranno automaticamente i dati prodotti localmente nella regione dall'argomento locale (ad esempio, `topic`), nonché i dati replicati dall'altra regione nell'argomento con il prefisso `<sourceKafkaClusterAlias>.topic`. Questi due set di consumatori devono avere ID di gruppi di consumatori diversi in modo che gli offset dei gruppi di consumatori non vengano sovrascritti quando il replicatore MSK li copia nell'altro cluster.

Risoluzione dei problemi relativi al replicatore MSK

Argomenti

- Lo stato del replicatore MSK passa da CREATING a FAILED
- Il replicatore MSK appare bloccato nello stato CREATING

- Il replicatore MSK non replica dati o replica soltanto dati parziali
- Gli offset dei messaggi nel cluster di destinazione sono diversi da quelli del cluster di origine
- MSK Replicator non sincronizza gli offset dei gruppi di consumatori oppure il gruppo di consumatori non esiste nel cluster di destinazione.
- La latenza di replica è elevata o continua ad aumentare

Le seguenti informazioni agevolano la risoluzione dei problemi che si potrebbero verificare con il replicatore MSK. Puoi anche pubblicare il problema in AWS re:Post.

Lo stato del replicatore MSK passa da CREATING a FAILED

Di seguito sono riportate alcune cause comuni degli errori di creazione del replicatore MSK.

1. Assicurati che i gruppi di sicurezza che hai fornito per la creazione del replicatore nella sezione del cluster di destinazione dispongano di regole in uscita per consentire il traffico verso i gruppi di sicurezza del cluster di destinazione. Inoltre, assicurati che i gruppi di sicurezza del cluster di destinazione dispongano di regole in entrata che consentano il traffico verso i gruppi di sicurezza che fornisci per la creazione del replicatore nella sezione del cluster di destinazione. Per informazioni, consulta Scelta del cluster di destinazione.
2. Se stai creando il replicatore per la replica tra regioni, verifica che per il cluster di origine sia attivata la connettività multi-VPC per il metodo di autenticazione Controllo degli accessi IAM. Per informazioni, consulta Connettività privata multi-VPC di Amazon MSK in un'unica regione. Verifica inoltre che la policy del cluster sia configurata sul cluster di origine in modo che il replicatore MSK possa connettersi a esso. Per informazioni, consulta Passaggio 1: preparazione del cluster di origine Amazon MSK.
3. Assicurati che il ruolo IAM fornito durante la creazione del replicatore MSK disponga delle autorizzazioni necessarie per leggere e scrivere nei cluster di origine e di destinazione. Inoltre, verifica che il ruolo IAM disponga delle autorizzazioni per scrivere sugli argomenti. Per informazioni, consultare Configurazione delle impostazioni e delle autorizzazioni del replicatore.
4. Assicurati che le ACL di rete non blocchino la connessione tra il replicatore MSK e i cluster di origine e di destinazione.
5. È possibile che i cluster di origine o di destinazione non fossero completamente disponibili quando il replicatore MSK ha tentato di connettersi a essi. Ciò potrebbe essere dovuto a un carico eccessivo, all'utilizzo del disco o della CPU, che impedisce al replicatore di connettersi ai broker. Risolvi il problema con i broker e prova di nuovo a creare il replicatore.

Dopo aver eseguito le convalide precedenti, crea nuovamente il replicatore MSK.

Il replicatore MSK appare bloccato nello stato CREATING

A volte la creazione del replicatore MSK può richiedere fino a 30 minuti. Attendi 30 minuti e controlla nuovamente lo stato del replicatore.

Il replicatore MSK non replica dati o replica soltanto dati parziali

Seguire questi passaggi per risolvere i problemi di replica dei dati.

1. Verificate che il vostro Replicator non stia riscontrando errori di autenticazione utilizzando la AuthError metrica fornita da MSK Replicator in. CloudWatch Se questo parametro è superiore a 0, verifica se la policy del ruolo IAM fornito per il replicatore è valida e che non siano impostate autorizzazioni di rifiuto per le autorizzazioni del cluster. In base alla dimensione clusterAlias, è possibile verificare se è il cluster di origine o quello di destinazione a presentare errori di autenticazione.
2. Verifica che i cluster di origine e di destinazione non presentino problemi. È possibile che il replicatore non sia in grado di connettersi al cluster di origine o di destinazione. Ciò può accadere a causa di un numero eccessivo di connessioni, di un disco a piena capacità o di un elevato utilizzo della CPU.
3. Verificate che i cluster di origine e di destinazione siano raggiungibili da MSK Replicator utilizzando la metrica in. KafkaClusterPingSuccessCount CloudWatch In base alla dimensione clusterAlias, è possibile verificare se è il cluster di origine o di destinazione a presentare errori di autenticazione. Se questo parametro è 0 o non ha un punto di dati, la connessione non è integra. È necessario verificare le autorizzazioni di rete e i ruoli IAM utilizzati dal replicatore MSK per connettersi ai cluster.
4. Verificate che il Replicator non stia riscontrando errori dovuti alla mancanza di autorizzazioni a livello di argomento utilizzando la metrica in. ReplicatorFailure CloudWatch Se questo parametro è superiore a 0, controlla il ruolo IAM che hai fornito per le autorizzazioni a livello di argomento.
5. Verifica che l'espressione regolare che hai fornito nell'elenco consentito durante la creazione del replicatore corrisponda ai nomi degli argomenti che desideri replicare. Inoltre, verifica che gli argomenti non vengano esclusi dalla replica a causa di un'espressione regolare nell'elenco degli argomenti non consentiti.
6. Si noti che il Replicator potrebbe impiegare fino a 30 secondi per rilevare e creare nuovi argomenti o partizioni di argomenti sul cluster di destinazione. Tutti i messaggi inviati all'argomento di

origine prima della creazione dell'argomento nel cluster di destinazione non verranno replicati se la posizione iniziale del replicatore è la più recente (impostazione predefinita). In alternativa, è possibile avviare la replica dal primo offset nelle partizioni degli argomenti del cluster di origine se si desidera replicare i messaggi esistenti sui propri argomenti nel cluster di destinazione. Per informazioni, consulta Configurazione delle impostazioni e delle autorizzazioni del replicatore.

Gli offset dei messaggi nel cluster di destinazione sono diversi da quelli del cluster di origine

Nell'ambito della replica dei dati, MSK Replicator consuma i messaggi dal cluster di origine e li produce nel cluster di destinazione. Ciò può portare a messaggi con offset diversi sui cluster di origine e di destinazione. Tuttavia, se è stata attivata la sincronizzazione degli offset dei gruppi di consumatori durante la creazione di Replicator, MSK Replicator tradurrà automaticamente gli offset durante la copia dei metadati in modo che, dopo il failover sul cluster di destinazione, gli utenti possano riprendere l'elaborazione da dove l'avevano interrotta nel cluster di origine.

MSK Replicator non sincronizza gli offset dei gruppi di consumatori oppure il gruppo di consumatori non esiste nel cluster di destinazione.

Segui questi passaggi per risolvere i problemi di replica dei metadati.

1. Verifica che la replica dei dati funzioni come previsto. In caso contrario, vediIl replicatore MSK non replica dati o replica soltanto dati parziali.
2. Verifica che l'espressione regolare che hai fornito nell'elenco consentito durante la creazione del Replicator corrisponda ai nomi dei gruppi di consumatori che desideri replicare. Inoltre, verificate che i gruppi di consumatori non vengano esclusi dalla replica a causa di un'espressione regolare nell'elenco degli utenti non autorizzati.
3. Verificate che MSK Replicator abbia creato l'argomento sul cluster di destinazione. Potrebbero essere necessari fino a 30 secondi prima che il Replicator rilevi e crei i nuovi argomenti o le partizioni degli argomenti sul cluster di destinazione. Tutti i messaggi inviati all'argomento di origine prima della creazione dell'argomento nel cluster di destinazione non verranno replicati se la posizione iniziale del replicatore è la più recente (impostazione predefinita). Se il gruppo di consumatori nel cluster di origine ha utilizzato solo i messaggi che non sono stati replicati da MSK Replicator, il gruppo di consumatori non verrà replicato nel cluster di destinazione. Dopo aver creato correttamente l'argomento sul cluster di destinazione, MSK Replicator inizierà a replicare i nuovi messaggi scritti dal cluster di origine alla destinazione. Una volta che il

gruppo di consumatori inizia a leggere questi messaggi dall'origine, MSK Replicator replicherà automaticamente il gruppo di consumatori nel cluster di destinazione. In alternativa, è possibile avviare la replica dal primo offset nelle partizioni degli argomenti del cluster di origine se si desidera replicare i messaggi esistenti sui propri argomenti nel cluster di destinazione. Per informazioni, consulta Configurazione delle impostazioni e delle autorizzazioni del replicatore.

> **Note**
>
> MSK Replicator ottimizza la sincronizzazione dell'offset dei gruppi di consumatori per i consumatori del cluster di origine che leggono da una posizione più vicina alla fine della partizione degli argomenti. Se i gruppi di consumatori sono in ritardo rispetto al cluster di origine, è possibile riscontrare un ritardo maggiore per tali gruppi di consumatori sul cluster di destinazione rispetto a quello di origine. Ciò significa che, dopo il failover sul cluster di destinazione, i consumatori rielaboreranno più messaggi duplicati. Per ridurre questo ritardo, i tuoi utenti del cluster di origine dovrebbero recuperare il ritardo e iniziare a consumare dall'estremità dello stream (fine della partizione dell'argomento). Man mano che i consumatori recuperano il ritardo, MSK Replicator ridurrà automaticamente il ritardo.

La latenza di replica è elevata o continua ad aumentare

Di seguito sono riportate alcune cause comuni dell'elevata latenza di replica.

1. Verifica di disporre del numero corretto di partizioni nei cluster MSK di origine e di destinazione. Un numero di partizioni troppo basso o elevato può influire sulle prestazioni. Per indicazioni sulla scelta del numero di partizioni, consulta la sezione Best practice per l'utilizzo del replicatore MSK. La tabella seguente mostra il numero minimo di partizioni consigliato per ottenere la velocità di trasmissione effettiva desiderata con il replicatore MSK.

   Velocità di trasmissione effettiva e numero minimo consigliato di partizioni

Velocità di trasmissione effettiva (MB/s)	Numero minimo di partizioni necessarie
50	167
100	334
250	833

Velocità di trasmissione effettiva (MB/s)	Numero minimo di partizioni necessarie
500	1666
1000	3333

2. Verifica di disporre di una capacità di lettura e scrittura sufficiente nei cluster MSK di origine e di destinazione per supportare il traffico di replica. Il replicatore MSK funge da consumatore per il cluster di origine (uscita) e da produttore per il cluster di destinazione (ingresso). Pertanto, è necessario fornire la capacità del cluster per supportare il traffico di replica oltre al resto del traffico sui cluster. Consulta la sezione ??? per indicazioni sul dimensionamento dei cluster MSK.
3. La latenza di replica può variare per i cluster MSK in diverse coppie di AWS regioni di origine e destinazione, a seconda della distanza geografica dei cluster l'uno dall'altro. Ad esempio, la latenza di replica è in genere inferiore quando si esegue la replica tra cluster nelle regioni Europa (Irlanda) ed Europa (Londra) rispetto alla replica tra cluster nelle regioni Europa (Irlanda) e Asia Pacifico (Sydney).
4. Assicurati che il replicatore non subisca limitazioni a causa delle quote eccessivamente aggressive impostate sui cluster di origine o di destinazione. È possibile utilizzare la ThrottleTime metrica fornita da MSK Replicator in per visualizzare il tempo medio, in millisecondi, in CloudWatch cui una richiesta è stata limitata dai broker del cluster di origine/destinazione. Se questo parametro è superiore a 0, è necessario modificare le quote Kafka per ridurre la limitazione della larghezza di banda della rete in modo che il replicatore possa recuperare il ritardo. Per informazioni sulla gestione delle quote Kafka per il replicatore, consulta la pagina Gestione della velocità di trasmissione effettiva del replicatore MSK utilizzando le quote Kafka.
5. ReplicationLatency e AWS potrebbe aumentare quando una regione si degrada. MessageLag Utilizza Dashboard AWS Service Health per verificare la presenza di un evento del servizio MSK nella regione in cui si trova il cluster MSK primario. Se si verifica un evento di servizio, è possibile reindirizzare temporaneamente le operazioni di lettura e scrittura dell'applicazione all'altra regione.

Best practice per l'utilizzo del replicatore MSK

Questa sezione descrive le best practice e le strategie di implementazione comuni per l'utilizzo del replicatore MSK.

Argomenti

- Gestione della velocità di trasmissione effettiva del replicatore MSK utilizzando le quote Kafka

- Impostazione del periodo di conservazione dei cluster

Gestione della velocità di trasmissione effettiva del replicatore MSK utilizzando le quote Kafka

Poiché il replicatore MSK funge da consumatore per il cluster di origine, la replica può causare la limitazione della larghezza di banda della rete di altri consumatori sul cluster di origine. L'entità della limitazione della larghezza di banda della rete dipende dalla capacità di lettura disponibile sul cluster di origine e dalla velocità di trasmissione effettiva dei dati da replicare. Ti consigliamo di fornire una capacità identica per i cluster di origine e di destinazione e di tenere conto della velocità di trasmissione effettiva di replica nel calcolo della capacità necessaria.

È inoltre possibile impostare quote Kafka per il replicatore sui cluster di origine e di destinazione per controllare la capacità che il replicatore MSK può utilizzare. Si consiglia di specificare una quota di larghezza di banda della rete. Una quota di larghezza di banda della rete definisce una soglia di velocità di byte, espressa in byte al secondo, per uno o più client che condividono una quota. Questa quota è definita per singolo broker.

Segui questi passaggi per applicare una quota.

1. Recupera la stringa del server di bootstrap per il cluster di origine. Per informazioni, consulta Recupero dei broker di bootstrap per un cluster Amazon MSK.
2. Recupera il ruolo di esecuzione del servizio (SER) utilizzato dal replicatore MSK. Questo è il SER che hai utilizzato per una richiesta `CreateReplicator`. È inoltre possibile estrarre il SER dalla DescribeReplicator risposta di un replicatore esistente.
3. Utilizzando gli strumenti della CLI di Kafka, esegui il seguente comando sul cluster di origine.

```
./kafka-configs.sh --bootstrap-server <source-cluster-bootstrap-server> --alter --
add-config 'consumer_byte_
rate=<quota_in_bytes_per_second>' --entity-type users --entity-name
 arn:aws:sts::<customer-account-id>:assumed-role/<ser-role-name>/<customer-account-
id> --command-config <client-properties-for-iam-auth></programlisting>
```

4. Dopo aver eseguito il comando precedente, verifica che il parametro `ReplicatorThroughput` non superi la quota impostata.

Nota che se riutilizzi un ruolo di esecuzione del servizio tra più replicatori MSK, questi sono tutti soggetti a questa quota. Se desideri mantenere quote separate per il replicatore, utilizza ruoli di esecuzione del servizio separati.

Per ulteriori informazioni sull'utilizzo dell'autenticazione IAM di MSK con le quote, consulta la pagina Multi-tenancy Apache Kafka clusters in Amazon MSK with IAM access control and Kafka Quotas – Part 1.

> **Warning**
>
> L'impostazione di un valore consumer_byte_rate estremamente basso può causare comportamenti inaspettati da parte del replicatore MSK.

Impostazione del periodo di conservazione dei cluster

È possibile impostare il periodo di conservazione dei log sia per i cluster MSK assegnati sia per quelli serverless. Il periodo di conservazione consigliato è di 7 giorni. Vedi Modifiche alla configurazione del cluster o Configurazione del cluster MSK Serverless.

Stati dei cluster

Nella tabella seguente vengono descritti gli stati possibili di un cluster e il rispettivo significato. Descrive inoltre quali operazioni è possibile e non è possibile eseguire quando un cluster si trova in uno di questi stati. Per scoprire lo stato di un cluster, consulta la AWS Management Console. È inoltre possibile utilizzare il comando describe-cluster-v2 o l'operazione DescribeClusterV2 per descrivere il cluster. La descrizione di un cluster include il relativo stato.

Stato del cluster	Significato e operazioni possibili
ACTIVE	Puoi produrre e utilizzare dati. Puoi anche eseguire l'API e AWS CLI le operazioni di Amazon MSK sul cluster.
CREAZIONE IN CORSO	Amazon MSK sta configurando il cluster. È necessario attendere che il cluster raggiunga lo stato ATTIVO prima di poterlo utilizzare per produrre o consumare dati o per eseguire l'API o AWS CLI le operazioni di Amazon MSK su di esso.
ELIMINAZIONE IN CORSO	Il cluster è in fase di eliminazione. Non è possibile utilizzarlo per produrre o utilizzare dati. Inoltre, non è possibile eseguire l'API o AWS CLI le operazioni di Amazon MSK su di essa.
Non riuscito	Il processo di creazione o eliminazione del cluster non è riuscito. Non è possibile utilizzar e il cluster per produrre o utilizzare dati. Puoi eliminare il cluster ma non puoi eseguire operazioni di API Amazon MSK o AWS CLI aggiornarlo.
HEALING	Amazon MSK sta eseguendo un'operazione interna, ad esempio la sostituzione di un broker non funzionante. Ad esempio, il broker

Stato del cluster	Significato e operazioni possibili
	potrebbe non rispondere. È ancora possibile utilizzare il cluster per produrre e utilizzare dati. Tuttavia, non è possibile eseguire l'API di Amazon MSK o AWS CLI aggiornare le operazioni sul cluster finché non torna allo stato ACTIVE.
MAINTENANCE	Amazon MSK esegue operazioni di manutenzione ordinaria sul cluster. Tali operazioni di manutenzione includono l'applicazione di patch di sicurezza. È ancora possibile utilizzare il cluster per produrre e utilizzare dati. Tuttavia, non è possibile eseguire l'API di Amazon MSK o AWS CLI aggiornare le operazioni sul cluster finché non torna allo stato ACTIVE.
REBOOTING_BROKER	Amazon MSK sta riavviando un broker. È ancora possibile utilizzare il cluster per produrre e utilizzare dati. Tuttavia, non è possibile eseguire l'API di Amazon MSK o AWS CLI aggiornare le operazioni sul cluster finché non torna allo stato ACTIVE.
AGGIORNAMENTO IN CORSO	Un'API o un' AWS CLI operazione Amazon MSK avviata dall'utente sta aggiornando il cluster. È ancora possibile utilizzare il cluster per produrre e utilizzare dati. Tuttavia, non è possibile eseguire alcuna API Amazon MSK aggiuntiva o eseguire operazioni di AWS CLI aggiornamento sul cluster finché non torna allo stato ATTIVO.

Sicurezza in Streaming gestito da Amazon per Apache Kafka

La sicurezza del cloud AWS è la massima priorità. In qualità di AWS cliente, puoi beneficiare di un data center e di un'architettura di rete progettati per soddisfare i requisiti delle organizzazioni più sensibili alla sicurezza.

La sicurezza è una responsabilità condivisa tra AWS te e te. Il modello di responsabilità condivisa descrive questo aspetto come sicurezza del cloud e sicurezza nel cloud:

- Sicurezza del cloud: AWS è responsabile della protezione dell'infrastruttura che gestisce AWS i servizi nel AWS cloud. AWS ti fornisce anche servizi che puoi utilizzare in modo sicuro. I revisori esterni testano e verificano regolarmente l'efficacia della nostra sicurezza nell'ambito dei AWS Programmi di AWS conformità dei Programmi di conformità dei di . Per ulteriori informazioni sui programmi di conformità applicabili a Streaming gestito da Amazon per Apache Kafka, consulta la pagina Amazon Web Services in Scope by Compliance Program.
- Sicurezza nel cloud: la responsabilità dell'utente è determinata dal AWS servizio che utilizza. Inoltre, sei responsabile anche di altri fattori, tra cui la riservatezza dei dati, i requisiti dell'azienda e le leggi e le normative applicabili.

La presente documentazione aiuta a comprendere come applicare il modello di responsabilità condivisa quando si utilizza Amazon MSK. Gli argomenti seguenti descrivono come configurare Amazon MSK per soddisfare gli obiettivi di sicurezza e conformità. Vengono inoltre fornite informazioni su come utilizzare altri servizi di Amazon Web Services che consentono di monitorare e proteggere le risorse Amazon MSK.

Argomenti

- Resilienza in Streaming gestito da Amazon per Apache Kafka
- Sicurezza dell'infrastruttura in Streaming gestito da Amazon per Apache Kafka

Protezione dei dati in Streaming gestito da Amazon per Apache Kafka

Il modello di responsabilità AWS condivisa modello si applica alla protezione dei dati in Amazon Managed Streaming for Apache Kafka. Come descritto in questo modello, AWS è responsabile della protezione dell'infrastruttura globale che gestisce tutti i. Cloud AWS L'utente è responsabile del controllo dei contenuti ospitati su questa infrastruttura. L'utente è inoltre responsabile della configurazione della protezione e delle attività di gestione per i Servizi AWS utilizzati. Per ulteriori informazioni sulla privacy dei dati, vedi le Domande frequenti sulla privacy dei dati. Per informazioni sulla protezione dei dati in Europa, consulta il post del blog relativo al Modello di responsabilità condivisa AWS e GDPR nel Blog sulla sicurezza AWS .

Ai fini della protezione dei dati, consigliamo di proteggere Account AWS le credenziali e configurare i singoli utenti con AWS IAM Identity Center or AWS Identity and Access Management (IAM). In tal modo, a ogni utente verranno assegnate solo le autorizzazioni necessarie per svolgere i suoi compiti. Ti suggeriamo, inoltre, di proteggere i dati nei seguenti modi:

- Utilizza l'autenticazione a più fattori (MFA) con ogni account.
- Usa SSL/TLS per comunicare con le risorse. AWS È richiesto TLS 1.2 ed è consigliato TLS 1.3.
- Configura l'API e la registrazione delle attività degli utenti con. AWS CloudTrail
- Utilizza soluzioni di AWS crittografia, insieme a tutti i controlli di sicurezza predefiniti all'interno Servizi AWS.
- Utilizza i servizi di sicurezza gestiti avanzati, come Amazon Macie, che aiutano a individuare e proteggere i dati sensibili archiviati in Amazon S3.
- Se hai bisogno di moduli crittografici convalidati FIPS 140-2 per l'accesso AWS tramite un'interfaccia a riga di comando o un'API, utilizza un endpoint FIPS. Per ulteriori informazioni sugli endpoint FIPS disponibili, consulta il Federal Information Processing Standard (FIPS) 140-2.

Ti consigliamo vivamente di non inserire mai informazioni riservate o sensibili, ad esempio gli indirizzi e-mail dei clienti, nei tag o nei campi di testo in formato libero, ad esempio nel campo Nome. Ciò include quando lavori con Amazon MSK o altro Servizi AWS utilizzando la console, l'API o gli AWS SDK. AWS CLI I dati inseriti nei tag o nei campi di testo in formato libero utilizzati per i nomi

possono essere utilizzati per i la fatturazione o i log di diagnostica. Quando fornisci un URL a un server esterno, ti suggeriamo vivamente di non includere informazioni sulle credenziali nell'URL per convalidare la tua richiesta al server.

Argomenti

Crittografia di Amazon MSK

Amazon MSK fornisce opzioni di crittografia dei dati che puoi utilizzare per soddisfare rigidi requisiti di gestione dei dati. I certificati utilizzati da Amazon MSK per la crittografia devono essere rinnovati ogni 13 mesi. Amazon MSK rinnova automaticamente questi certificati per tutti i cluster. Imposta lo stato del cluster su `MAINTENANCE` quando avvia l'operazione di aggiornamento del certificato. Lo stato viene reimpostato su `ACTIVE` al termine dell'aggiornamento. Mentre un cluster è nello stato `MAINTENANCE`, è possibile continuare a produrre e consumare dati, ma non è possibile eseguire alcuna operazione di aggiornamento su di esso.

Crittografia a riposo

Amazon MSK si integra con AWS Key Management Service (KMS) per offrire una crittografia lato server trasparente. Amazon MSK esegue sempre la crittografia dei dati a riposo. Quando crei un cluster Amazon MSK, puoi specificare la AWS KMS key che desideri far utilizzare ad Amazon MSK per crittografare i dati a riposo. Se non specifichi una chiave KMS, Amazon MSK crea una chiave KMS gestita da Chiave gestita da AWS e la utilizza per tuo conto. Per ulteriori informazioni sulle chiavi KMS, consultare AWS KMS keys nella Guida per sviluppatori di AWS Key Management Service .

Crittografia in transito

Amazon MSK utilizza TLS 1.2. Per impostazione predefinita, effettua la crittografia dei dati in transito tra i broker del cluster MSK. Puoi ignorare questa impostazione predefinita al momento della creazione del cluster.

Per la comunicazione tra client e broker, è necessario specificare una delle tre impostazioni seguenti:

- Consenti solo dati crittografati TLS. Si tratta dell'impostazione di default.
- Consenti dati non crittografati e dati crittografati TLS.

- Consenti solo dati non crittografati.

I broker Amazon MSK utilizzano certificati pubblici AWS Certificate Manager . Pertanto, qualsiasi truststore che considera attendibili gli Amazon Trust Services considera attendibili anche i certificati dei broker Amazon MSK.

Anche se consigliamo vivamente di abilitare la crittografia dei dati in transito, questa potrebbe aggiungere un sovraccarico aggiuntivo della CPU e alcuni millisecondi di latenza. Tuttavia, la maggior parte dei casi d'uso non è sensibile a queste differenze e l'entità dell'impatto dipende dalla configurazione del cluster, dei client e del profilo di utilizzo.

Quali sono i primi passi per iniziare a utilizzare la crittografia?

Durante la creazione di un cluster MSK, puoi specificare le impostazioni di crittografia in formato JSON. Di seguito è riportato un esempio.

```
{
   "EncryptionAtRest": {
       "DataVolumeKMSKeyId": "arn:aws:kms:us-east-1:123456789012:key/abcdabcd-1234-
abcd-1234-abcd123e8e8e"
    },
   "EncryptionInTransit": {
        "InCluster": true,
        "ClientBroker": "TLS"
    }
}
```

Per `DataVolumeKMSKeyId`, puoi specificare una chiave gestita dal cliente o la Chiave gestita da AWS per MSK nel tuo account (`alias/aws/kafka`). Se non lo specifichi`EncryptionAtRest`, Amazon MSK crittografa comunque i tuoi dati inattivi in base a. Chiave gestita da AWS Per determinare la chiave utilizzata dal cluster, invia una richiesta `GET` o richiama l'operazione API `DescribeCluster`.

Per `EncryptionInTransit`, il valore predefinito di `InCluster` è true, ma puoi impostarlo su false se Amazon MSK non deve crittografare i dati durante il passaggio tra i broker.

Per specificare la modalità di crittografia per i dati in transito tra client e broker, imposta `ClientBroker` su uno di tre valori: `TLS`, `TLS_PLAINTEXT` o `PLAINTEXT`.

Per specificare le impostazioni di crittografia durante la creazione di un cluster

1. Salvare il contenuto dell'esempio precedente in un file e assegnare al file il nome desiderato. Ad esempio, chiamarlo `encryption-settings.json`.
2. Eseguire il comando `create-cluster` e utilizzare l'opzione `encryption-info` per puntare al file in cui il JSON di configurazione è stato salvato. Di seguito è riportato un esempio. Sostituisci *`{YOUR MSK VERSION}`* con una versione che corrisponda alla versione del client Apache Kafka. Per informazioni su come trovare la versione del cluster MSK in uso, consulta la pagina To find the version of your MSK cluster. Tieni presente che l'utilizzo di una versione del client Apache Kafka diversa da quella del cluster MSK può causare il danneggiamento, la perdita dei dati e tempi di inattività di Apache Kafka.

```
aws kafka create-cluster --cluster-name "ExampleClusterName" --broker-node-group-
info file://brokernodegroupinfo.json --encryption-info file://encryptioninfo.json
 --kafka-version "{YOUR MSK VERSION}" --number-of-broker-nodes 3
```

Di seguito è riportato un esempio di una risposta corretta dopo l'esecuzione di questo comando.

```
{
    "ClusterArn": "arn:aws:kafka:us-east-1:123456789012:cluster/SecondTLSTest/
abcdabcd-1234-abcd-1234-abcd123e8e8e",
    "ClusterName": "ExampleClusterName",
    "State": "CREATING"
}
```

Per testare la crittografia TLS

1. Creare un computer client seguendo le linee guida in the section called “Passaggio 3: creazione di un computer client”.
2. Installare Apache Kafka sul computer client.
3. In questo esempio, utilizziamo il truststore JVM per comunicare con il cluster MSK. A tale scopo, creare innanzitutto una cartella denominata `/tmp` sul computer client. Quindi, passare alla cartella `bin` dell'installazione di Apache Kafka ed eseguire il seguente comando. (Il percorso JVM potrebbe essere diverso.)

```
cp /usr/lib/jvm/java-1.8.0-openjdk-1.8.0.201.b09-0.amzn2.x86_64/jre/lib/security/
cacerts /tmp/kafka.client.truststore.jks
```

4. Sempre nella cartella `bin` dell'installazione di Apache Kafka sul computer client, creare un file di testo denominato `client.properties` con il seguente contenuto.

```
security.protocol=SSL
ssl.truststore.location=/tmp/kafka.client.truststore.jks
```

5. Esegui il comando seguente su un computer su cui è AWS CLI installato, sostituendo *`clusterARN`* con l'ARN del tuo cluster.

```
aws kafka get-bootstrap-brokers --cluster-arn clusterARN
```

Se l'operazione riesce, il risultato sarà simile al seguente. Salvare questo risultato perché è necessario per la fase successiva.

```
{
    "BootstrapBrokerStringTls": "a-1.example.g7oein.c2.kafka.us-
east-1.amazonaws.com:0123,a-3.example.g7oein.c2.kafka.us-
east-1.amazonaws.com:0123,a-2.example.g7oein.c2.kafka.us-east-1.amazonaws.com:0123"
}
```

6. Esegui il comando seguente, sostituendolo *`BootstrapBrokerStringTls`*con uno degli endpoint del broker ottenuti nel passaggio precedente.

```
<path-to-your-kafka-installation>/bin/kafka-console-producer.sh --broker-
list BootstrapBrokerStringTls --producer.config client.properties --topic
 TLSTestTopic
```

7. Apri una nuova finestra di comando e connettiti allo stesso computer client. Quindi, esegui il comando seguente per creare un utente della console.

```
<path-to-your-kafka-installation>/bin/kafka-console-consumer.sh --bootstrap-
server BootstrapBrokerStringTls --consumer.config client.properties --topic
 TLSTestTopic
```

8. Nella finestra del produttore, digita un messaggio di testo seguito da un invio e cerca lo stesso messaggio nella finestra del consumatore. Amazon MSK ha crittografato questo messaggio in transito.

Per ulteriori informazioni sulla configurazione dei client Apache Kafka per l'utilizzo di dati crittografati, consulta Configuring Kafka Clients.

Autenticazione e autorizzazione per le API di Amazon MSK

AWS Identity and Access Management (IAM) è un software Servizio AWS che aiuta un amministratore a controllare in modo sicuro l'accesso alle AWS risorse. Gli amministratori IAM controllano chi può essere autenticato (accesso effettuato) e autorizzato (dotato di autorizzazioni) per utilizzare le risorse Amazon MSK. IAM è un software Servizio AWS che puoi utilizzare senza costi aggiuntivi.

Questa pagina descrive come utilizzare IAM per controllare chi può eseguire le operazioni di Amazon MSK sul tuo cluster. Per informazioni su come controllare chi può eseguire le operazioni di Apache Kafka sul tuo cluster, consulta la pagina the section called “Autenticazione e autorizzazione per le API di Apache Kafka”.

Argomenti

- Funzionamento di Amazon MSK con IAM
- Esempi di policy basate sull'identità per Amazon MSK
- Utilizzo di ruoli collegati ai servizi per Amazon MSK
- AWS politiche gestite per Amazon MSK
- Risoluzione dei problemi relativi all'identità e all'accesso di Amazon MSK

Funzionamento di Amazon MSK con IAM

Prima di utilizzare IAM per gestire l'accesso ad Amazon MSK, è necessario comprendere quali funzioni IAM sono disponibili per l'uso con Amazon MSK. Per avere una visione di alto livello di come Amazon MSK e altri AWS servizi funzionano con IAM, consulta AWS Services That Work with IAM nella IAM User Guide.

Argomenti

- Policy basate sull'identità di Amazon MSK
- Policy basate sulle risorse di Amazon MSK
- AWS politiche gestite
- Autorizzazione basata sui tag Amazon MSK

- Ruoli IAM di Amazon MSK

Policy basate sull'identità di Amazon MSK

Con le policy basate su identità di IAM, è possibile specificare quali azioni e risorse sono consentite o rifiutate, nonché le condizioni in base alle quali le azioni sono consentite o rifiutate. Amazon MSK supporta operazioni, risorse e chiavi di condizione specifiche. Per informazioni su tutti gli elementi utilizzati in una policy JSON, consulta Documentazione di riferimento degli elementi delle policy JSON IAM nella Guida per l'utente IAM.

Azioni

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. Cioè, quale principale può eseguire azioni su quali risorse, e in quali condizioni.

L'elemento `Action`di una policy JSON descrive le azioni che è possibile utilizzare per consentire o negare l'accesso a un criterio. Le azioni politiche in genere hanno lo stesso nome dell'operazione AWS API associata. Ci sono alcune eccezioni, ad esempio le azioni di sola autorizzazione che non hanno un'operazione API corrispondente. Esistono anche alcune operazioni che richiedono più operazioni in una policy. Queste operazioni aggiuntive sono denominate operazioni dipendenti.

Includi le operazioni in una policy per concedere le autorizzazioni a eseguire l'operazione associata.

Le operazioni delle policy in Amazon MSK utilizzano il seguente prefisso prima dell'operazione: `kafka:`. Ad esempio, per concedere a qualcuno l'autorizzazione per descrivere un cluster MSK con l'operazione API `DescribeCluster` di Amazon MSK, includi l'operazione `kafka:DescribeCluster` nella policy. Le istruzioni della policy devono includere un elemento `Action` o `NotAction`. Amazon MSK definisce un proprio set di operazioni che descrivono le attività che puoi eseguire con quel servizio.

Per specificare più azioni in una sola istruzione, separa ciascuna di esse con una virgola come mostrato di seguito:

```
"Action": ["kafka:action1", "kafka:action2"]
```

È possibile specificare più azioni tramite caratteri jolly (*). Ad esempio, per specificare tutte le azioni che iniziano con la parola `Describe`, includi la seguente azione:

```
"Action": "kafka:Describe*"
```

Per visualizzare un elenco delle operazioni di Amazon MSK, consulta la pagina Actions, resources, and condition keys for Amazon Managed Streaming for Apache Kafka nella Guida per l'utente di IAM.

Risorse

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. Cioè, quale principale può eseguire operazioni su quali risorse, e in quali condizioni.

L'elemento JSON `Resource`della policy specifica l'oggetto o gli oggetti ai quali si applica l'operazione. Le istruzioni devono includere un elemento `Resource`o un elemento `NotResource`. Come best practice, specifica una risorsa utilizzando il suo nome della risorsa Amazon (ARN). Puoi eseguire questa operazione per azioni che supportano un tipo di risorsa specifico, note come autorizzazioni a livello di risorsa.

Per le azioni che non supportano le autorizzazioni a livello di risorsa, ad esempio le operazioni di elenco, utilizza un carattere jolly (*) per indicare che l'istruzione si applica a tutte le risorse.

```
"Resource": "*"
```

La risorsa istanza di Amazon MSK dispone del seguente ARN:

```
arn:${Partition}:kafka:${Region}:${Account}:cluster/${ClusterName}/${UUID}
```

Per ulteriori informazioni sul formato degli ARN, consulta Amazon Resource Names (ARNs) e AWS Service Namespaces.

Ad esempio, per specificare l'istanza `CustomerMessages` nell'istruzione, utilizza il seguente ARN:

```
"Resource": "arn:aws:kafka:us-east-1:123456789012:cluster/CustomerMessages/abcd1234-
abcd-dcba-4321-a1b2abcd9f9f-2"
```

Per specificare tutti le istanze che appartengono ad un account specifico, utilizza il carattere jolly (*):

```
"Resource": "arn:aws:kafka:us-east-1:123456789012:cluster/*"
```

Alcune operazioni di Amazon MSK, ad esempio quelle per la creazione delle risorse, non possono essere eseguite su una risorsa specifica. In questi casi, è necessario utilizzare il carattere jolly (*).

```
"Resource": "*"
```

Per specificare più risorse in una singola istruzione, separa gli ARN con le virgole.

```
"Resource": ["resource1", "resource2"]
```

Per visualizzare un elenco dei tipi di risorse Amazon MSK e dei relativi ARN, consulta la pagina Resources Defined by Amazon Managed Streaming for Apache Kafka nella Guida per l'utente di IAM. Per informazioni sulle operazioni con cui è possibile specificare l'ARN di ogni risorsa, consulta la pagina Actions Defined by Amazon Managed Service for Apache Kafka.

Chiavi di condizione

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. Cioè, quale principale può eseguire azioni su quali risorse, e in quali condizioni.

L'elemento `Condition`(o blocco `Condition`) consente di specificare le condizioni in cui un'istruzione è in vigore. L'elemento `Condition`è facoltativo. Puoi compilare espressioni condizionali che utilizzano operatori di condizione, ad esempio uguale a o minore di, per soddisfare la condizione nella policy con i valori nella richiesta.

Se specifichi più elementi `Condition`in un'istruzione o più chiavi in un singolo elemento `Condition`, questi vengono valutati da AWS utilizzando un'operazione `AND`logica. Se si specificano più valori per una singola chiave di condizione, AWS valuta la condizione utilizzando un'operazione logica. `OR` Tutte le condizioni devono essere soddisfatte prima che le autorizzazioni dell'istruzione vengano concesse.

Puoi anche utilizzare variabili segnaposto quando specifichi le condizioni. Ad esempio, puoi autorizzare un utente IAM ad accedere a una risorsa solo se è stata taggata con il relativo nome utente IAM. Per ulteriori informazioni, consulta Elementi delle policy IAM: variabili e tag nella Guida per l'utente di IAM.

AWS supporta chiavi di condizione globali e chiavi di condizione specifiche del servizio. Per visualizzare tutte le chiavi di condizione AWS globali, consulta le chiavi di contesto delle condizioni AWS globali nella Guida per l'utente IAM.

Amazon MSK definisce il proprio set di chiavi di condizione e supporta anche l'utilizzo di alcune chiavi di condizione globali. Per vedere tutte le chiavi di condizione AWS globali, consulta AWS Global Condition Context Keys nella Guida per l'utente IAM.

Per visualizzare un elenco delle chiavi di condizione di Amazon MSK, consulta la pagina Condition Keys for Amazon Managed Streaming for Apache Kafka nella Guida per l'utente di IAM. Per informazioni sulle operazioni e le risorse con le quali è possibile utilizzare una chiave di condizione, consulta la pagina Actions Defined by Amazon Managed Service for Apache Kafka.

Esempi

Per visualizzare degli esempi di policy basate sull'identità di Amazon MSK, consulta la pagina Esempi di policy basate sull'identità per Amazon MSK.

Policy basate sulle risorse di Amazon MSK

Amazon MSK supporta una policy del cluster (nota anche come policy basata sulle risorse) da utilizzare con i cluster Amazon MSK. Puoi utilizzare una policy del cluster per definire quali principali IAM dispongono delle autorizzazioni multi-account per configurare la connettività privata al tuo cluster Amazon MSK. In combinazione con l'autenticazione del client IAM, puoi utilizzare la policy del cluster anche per definire in modo granulare le autorizzazioni del piano dati Kafka per i client che si connettono.

Per visualizzare un esempio di come configurare una policy del cluster, consulta la sezione Passaggio 2: collegamento di una policy del cluster al cluster MSK.

AWS politiche gestite

Autorizzazione basata sui tag Amazon MSK

È possibile associare tag ai cluster Amazon MSK. Per controllare l'accesso basato su tag, fornisci informazioni sui tag nell'elemento condizione di una policy utilizzando le chiavi di condizione `kafka:ResourceTag/`*`key-name`*, `aws:RequestTag/`*`key-name`* o `aws:TagKeys`. Per ulteriori informazioni sull'assegnazione di tag alle risorse Amazon MSK, consulta la pagina the section called “Assegnazione di tag a un cluster”.

Per visualizzare una policy basata sulle identità di esempio per limitare l'accesso a una risorsa basata sui tag in tale cluster, consulta Accesso ai cluster Amazon MSK in base ai tag.

Ruoli IAM di Amazon MSK

Un ruolo IAM è un'entità all'interno dell'account Amazon Web Services che dispone di autorizzazioni specifiche.

Utilizzo delle credenziali temporanee con Amazon MSK

È possibile utilizzare credenziali temporanee per effettuare l'accesso con la federazione, assumere un ruolo IAM o un ruolo multi-account. È possibile ottenere credenziali di sicurezza temporanee chiamando operazioni AWS STS API come AssumeRoleo GetFederationToken.

Amazon MSK supporta l'uso delle credenziali temporanee.

Ruoli collegati ai servizi

I ruoli collegati ai servizi permettono ad Amazon Web Services di accedere a risorse in altri servizi per completare un'operazione a tuo nome. I ruoli collegati ai servizi sono visualizzati nell'account IAM e sono di proprietà del servizio. Un amministratore può visualizzare, ma non modificare le autorizzazioni dei ruoli collegati ai servizi.

Amazon MSK supporta i ruoli collegati ai servizi. Per maggiori dettagli su come creare e gestire i ruoli collegati ai servizi di Amazon MSK, consulta la pagina the section called “Ruoli collegati ai servizi”.

Esempi di policy basate sull'identità per Amazon MSK

Per impostazione predefinita, gli utenti e i ruoli IAM non sono autorizzati a eseguire le operazioni API di Amazon MSK. Un amministratore deve creare le policy IAM che concedono a utenti e ruoli l'autorizzazione per eseguire operazioni API specifiche sulle risorse specificate di cui hanno bisogno. L'amministratore deve quindi allegare queste policy a utenti o IAM che richiedono tali autorizzazioni.

Per informazioni su come creare una policy basata su identità IAM utilizzando questi documenti di policy JSON di esempio, consulta Creazione di policy nella scheda JSON nella Guida per l'utente IAM.

Argomenti

- Best practice delle policy
- Consentire agli utenti di visualizzare le loro autorizzazioni
- Accesso a un cluster Amazon MSK
- Accesso ai cluster Amazon MSK in base ai tag

Best practice delle policy

Le policy basate sull'identità determinano se qualcuno può creare, accedere o eliminare risorse Amazon MSK all'interno dell'account. Queste azioni possono comportare costi aggiuntivi per

l' Account AWS. Quando crei o modifichi policy basate su identità, segui queste linee guida e raccomandazioni:

- Inizia con le policy AWS gestite e passa alle autorizzazioni con privilegi minimi: per iniziare a concedere autorizzazioni a utenti e carichi di lavoro, utilizza le politiche AWS gestite che concedono le autorizzazioni per molti casi d'uso comuni. Sono disponibili nel tuo. Account AWS Ti consigliamo di ridurre ulteriormente le autorizzazioni definendo politiche gestite dai AWS clienti specifiche per i tuoi casi d'uso. Per ulteriori informazioni, consulta Policy gestite da AWSo Policy gestite da AWS per le funzioni dei processi nella Guida per l'utente IAM.
- Applica le autorizzazioni con privilegio minimo: quando imposti le autorizzazioni con le policy IAM, concedi solo le autorizzazioni richieste per eseguire un'attività. Puoi farlo definendo le azioni che possono essere intraprese su risorse specifiche in condizioni specifiche, note anche come autorizzazioni con privilegi minimi. Per ulteriori informazioni sull'utilizzo di IAM per applicare le autorizzazioni, consulta Policy e autorizzazioni in IAM nella Guida per l'utente IAM.
- Condizioni d'uso nelle policy IAM per limitare ulteriormente l'accesso: per limitare l'accesso a operazioni e risorse puoi aggiungere una condizione alle tue policy. Ad esempio, è possibile scrivere una condizione di policy per specificare che tutte le richieste devono essere inviate utilizzando SSL. Puoi anche utilizzare le condizioni per concedere l'accesso alle azioni del servizio se vengono utilizzate tramite uno specifico Servizio AWS, ad esempio AWS CloudFormation. Per ulteriori informazioni, consulta la sezione Elementi delle policy JSON di IAM: condizione nella Guida per l'utente IAM.
- Utilizzo di IAM Access Analyzer per convalidare le policy IAM e garantire autorizzazioni sicure e funzionali: IAM Access Analyzer convalida le policy nuove ed esistenti in modo che aderiscano alla sintassi della policy IAM (JSON) e alle best practice di IAM. IAM Access Analyzer offre oltre 100 controlli delle policy e consigli utili per creare policy sicure e funzionali. Per ulteriori informazioni, consulta Convalida delle policy per IAM Access Analyzer nella Guida per l'utente IAM.
- Richiedi l'autenticazione a più fattori (MFA): se hai uno scenario che richiede utenti IAM o un utente root nel Account AWS tuo, attiva l'MFA per una maggiore sicurezza. Per richiedere la MFA quando vengono chiamate le operazioni API, aggiungi le condizioni MFA alle policy. Per ulteriori informazioni, consulta Configurazione dell'accesso alle API protetto con MFA nella Guida per l'utente IAM.

Per maggiori informazioni sulle best practice in IAM, consulta Best practice di sicurezza in IAM nella Guida per l'utente di IAM.

Consentire agli utenti di visualizzare le loro autorizzazioni

Questo esempio mostra in che modo è possibile creare una policy che consente agli utenti IAM di visualizzare le policy inline e gestite che sono cpllegate alla relativa identità utente. Questa policy include le autorizzazioni per completare questa azione sulla console o utilizzando programmaticamente l'API o. AWS CLI AWS

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "ViewOwnUserInfo",
            "Effect": "Allow",
            "Action": [
                "iam:GetUserPolicy",
                "iam:ListGroupsForUser",
                "iam:ListAttachedUserPolicies",
                "iam:ListUserPolicies",
                "iam:GetUser"
            ],
            "Resource": ["arn:aws:iam::*:user/${aws:username}"]
        },
        {
            "Sid": "NavigateInConsole",
            "Effect": "Allow",
            "Action": [
                "iam:GetGroupPolicy",
                "iam:GetPolicyVersion",
                "iam:GetPolicy",
                "iam:ListAttachedGroupPolicies",
                "iam:ListGroupPolicies",
                "iam:ListPolicyVersions",
                "iam:ListPolicies",
                "iam:ListUsers"
            ],
            "Resource": "*"
        }
    ]
}
```

Accesso a un cluster Amazon MSK

In questo esempio, si desidera concedere a un utente IAM nell'account Amazon Web Services l'accesso a uno dei cluster, `purchaseQueriesCluster`. Questa policy consente all'utente di descrivere il cluster, ottenere i broker bootstrap, elencare i nodi broker e aggiornarlo.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Sid":"UpdateCluster",
         "Effect":"Allow",
         "Action":[
            "kafka:Describe*",
            "kafka:Get*",
            "kafka:List*",
            "kafka:Update*"
         ],
         "Resource":"arn:aws:kafka:us-east-1:012345678012:cluster/
purchaseQueriesCluster/abcdefab-1234-abcd-5678-cdef0123ab01-2"
      }
   ]
}
```

Accesso ai cluster Amazon MSK in base ai tag

Nella policy basata sull'identità, puoi utilizzare le condizioni per controllare l'accesso alle risorse Amazon MSK in base ai tag. In questo esempio viene illustrato come creare una policy che consente all'utente di descrivere il cluster, ottenere i broker bootstrap, elencare i nodi broker, aggiornarlo ed eliminarlo. Tuttavia, l'autorizzazione viene concessa solo se il valore del tag di cluster `0wner` è quello del nome utente.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AccessClusterIfOwner",
      "Effect": "Allow",
      "Action": [
        "kafka:Describe*",
        "kafka:Get*",
        "kafka:List*",
```

```
        "kafka:Update*",
        "kafka:Delete*"
      ],
      "Resource": "arn:aws:kafka:us-east-1:012345678012:cluster/*",
      "Condition": {
        "StringEquals": {
          "aws:ResourceTag/Owner": "${aws:username}"
        }
      }
    }
  ]
}
```

Puoi allegare questa policy agli utenti IAM nel tuo account. Se un utente denominato `richard-roe` tenta di aggiornare un cluster MSK, al cluster deve essere applicato il tag `Owner=richard-roe` o `owner=richard-roe`. In caso contrario, gli viene negato l'accesso. La chiave di tag di condizione `Owner` corrisponde a `Owner` e `owner` perché i nomi delle chiavi di condizione non effettuano la distinzione tra maiuscole e minuscole. Per ulteriori informazioni, consulta la sezione Elementi delle policy JSON di IAM: condizione nella Guida per l'utente di IAM.

Utilizzo di ruoli collegati ai servizi per Amazon MSK

Amazon MSK utilizza ruoli collegati ai servizi AWS Identity and Access Management (IAM). Un ruolo collegato ai servizi è un tipo di ruolo IAM univoco collegato direttamente ad Amazon MSK. I ruoli collegati ai servizi sono predefiniti da Amazon MSK e includono tutte le autorizzazioni richieste dal servizio per chiamare altri AWS servizi per tuo conto.

Un ruolo collegato ai servizi semplifica la configurazione di Amazon MSK perché consente di evitare l'aggiunta manuale delle autorizzazioni necessarie. Amazon MSK definisce le autorizzazioni dei ruoli collegati ai servizi. Se non diversamente definito, solo Amazon MSK può assumere i suoi ruoli. Le autorizzazioni definite includono la policy di attendibilità e la policy delle autorizzazioni che non può essere collegata a nessun'altra entità IAM.

Per informazioni sugli altri servizi che supportano i ruoli collegati ai servizi, consulta la pagina Amazon Web Services That Work with IAM e cerca i servizi che riportano Sì nella colonna Ruolo collegato ai servizi. Scegli Sì in corrispondenza di un link per visualizzare la documentazione relativa al ruolo collegato ai servizi per tale servizio.

Argomenti

- Autorizzazioni del ruolo collegato ai servizi per Amazon MSK

- Creazione di un ruolo collegato ai servizi per Amazon MSK
- Modifica di un ruolo collegato ai servizi per Amazon MSK
- Regioni supportate per i ruoli collegati ai servizi di Amazon MSK

Autorizzazioni del ruolo collegato ai servizi per Amazon MSK

Amazon MSK usa il ruolo collegato ai servizi denominato AWSServiceRoleForKafka. Amazon MSK utilizza questo ruolo per accedere alle risorse ed eseguire operazioni come:

- `*NetworkInterface`: crea e gestisci interfacce di rete nell'account cliente che rendano i broker del cluster accessibili ai client nel VPC del cliente.
- `*VpcEndpoints`— gestire gli endpoint VPC nell'account cliente che rendono i broker di cluster accessibili ai clienti nel VPC del cliente utilizzando. AWS PrivateLink Amazon MSK utilizza le autorizzazioni per `DescribeVpcEndpoints`, `ModifyVpcEndpoint` e `DeleteVpcEndpoints`.
- `secretsmanager`— gestisci le credenziali dei clienti con. AWS Secrets Manager
- `GetCertificateAuthorityCertificate`: recupera il certificato per la tua autorità di certificazione privata.

Questo ruolo collegato ai servizi è collegato alle seguenti policy gestite: `KafkaServiceRolePolicy`. Per gli aggiornamenti a questa politica, vedere KafkaServiceRolePolicy.

Ai fini dell'assunzione del ruolo, il ruolo collegato ai servizi AWSServiceRoleForKafka considera attendibili i seguenti servizi:

- `kafka.amazonaws.com`

La policy delle autorizzazioni del ruolo consente ad Amazon MSK di completare le seguenti operazioni sulle risorse.

```
{
 "Version": "2012-10-17",
 "Statement": [
  {
   "Effect": "Allow",
   "Action": [
    "ec2:CreateNetworkInterface",
```

```
    "ec2:DescribeNetworkInterfaces",
    "ec2:CreateNetworkInterfacePermission",
    "ec2:AttachNetworkInterface",
    "ec2:DeleteNetworkInterface",
    "ec2:DetachNetworkInterface",
    "ec2:DescribeVpcEndpoints",
    "acm-pca:GetCertificateAuthorityCertificate",
    "secretsmanager:ListSecrets"
   ],
   "Resource": "*"
  },
  {
   "Effect": "Allow",
   "Action": [
    "ec2:ModifyVpcEndpoint"
   ],
   "Resource": "arn:*:ec2:*:*:subnet/*"
  },
  {
   "Effect": "Allow",
   "Action": [
    "ec2:DeleteVpcEndpoints",
    "ec2:ModifyVpcEndpoint"
   ],
   "Resource": "arn:*:ec2:*:*:vpc-endpoint/*",
   "Condition": {
    "StringEquals": {
     "ec2:ResourceTag/AWSMSKManaged": "true"
    },
    "StringLike": {
     "ec2:ResourceTag/ClusterArn": "*"
    }
   }
  },
  {
   "Effect": "Allow",
   "Action": [
    "secretsmanager:GetResourcePolicy",
    "secretsmanager:PutResourcePolicy",
    "secretsmanager:DeleteResourcePolicy",
    "secretsmanager:DescribeSecret"
   ],
   "Resource": "*",
   "Condition": {
```

```
    "ArnLike": {
     "secretsmanager:SecretId": "arn:*:secretsmanager:*:*:secret:AmazonMSK_*"
    }
   }
  }
 ]
}
```

Per consentire a un'entità IAM (come un utente, un gruppo o un ruolo) di creare, modificare o eliminare un ruolo collegato ai servizi devi configurare le relative autorizzazioni. Per ulteriori informazioni, consulta Autorizzazioni del ruolo collegato ai servizi nella Guida per l'utente di IAM.

Creazione di un ruolo collegato ai servizi per Amazon MSK

Non è necessario creare manualmente un ruolo collegato ai servizi. Quando crei un cluster Amazon MSK nell'API AWS Management Console, nell'o nell' AWS API AWS CLI, Amazon MSK crea automaticamente il ruolo collegato al servizio.

Se elimini questo ruolo collegato ai servizi, puoi ricrearlo seguendo lo stesso processo utilizzato per ricreare il ruolo nell'account. Quando si crea un cluster Amazon MSK, Amazon MSK crea di nuovo automaticamente il ruolo collegato ai servizi per conto dell'utente.

Modifica di un ruolo collegato ai servizi per Amazon MSK

Amazon MSK non consente di modificare il ruolo collegato al servizio AWSServiceRoleForKafka. Dopo aver creato un ruolo collegato al servizio, non potrai modificarne il nome perché varie entità potrebbero farvi riferimento. È possibile tuttavia modificarne la descrizione utilizzando IAM. Per ulteriori informazioni, consulta la sezione Modifica di un ruolo collegato ai servizi nella Guida per l'utente di IAM.

Regioni supportate per i ruoli collegati ai servizi di Amazon MSK

Amazon MSK supporta l'utilizzo di ruoli collegati ai servizi in tutte le regioni in cui il servizio è disponibile. Per ulteriori informazioni, consulta AWS Regioni ed endpoint di .

AWS politiche gestite per Amazon MSK

Una politica AWS gestita è una politica autonoma creata e amministrata da. AWS AWS le politiche gestite sono progettate per fornire autorizzazioni per molti casi d'uso comuni, in modo da poter iniziare ad assegnare autorizzazioni a utenti, gruppi e ruoli.

Tieni presente che le policy AWS gestite potrebbero non concedere le autorizzazioni con il privilegio minimo per i tuoi casi d'uso specifici, poiché sono disponibili per tutti i clienti. AWS Ti consigliamo pertanto di ridurre ulteriormente le autorizzazioni definendo policy gestite dal cliente specifiche per i tuoi casi d'uso.

Non è possibile modificare le autorizzazioni definite nelle politiche gestite. AWS Se AWS aggiorna le autorizzazioni definite in una politica AWS gestita, l'aggiornamento ha effetto su tutte le identità principali (utenti, gruppi e ruoli) a cui è associata la politica. AWS è più probabile che aggiorni una policy AWS gestita quando ne Servizio AWS viene lanciata una nuova o quando diventano disponibili nuove operazioni API per i servizi esistenti.

Per ulteriori informazioni, consultare Policy gestite da AWSnella Guida per l'utente di IAM.

AWS politica gestita: AmazonMSK FullAccess

Questa policy concede autorizzazioni amministrative che consentono a un principale l'accesso completo a tutte le operazioni di Amazon MSK. Le autorizzazioni in questa policy sono raggruppate come segue:

- Le autorizzazioni Amazon MSK consentono tutte le operazioni di Amazon MSK.
- **`Amazon EC2`**autorizzazioni: in questa politica sono necessarie per convalidare le risorse passate in una richiesta API. Questo serve a garantire che Amazon MSK sia in grado di utilizzare correttamente le risorse di un cluster. Le altre autorizzazioni di Amazon EC2 incluse in questa policy consentono ad Amazon MSK di creare AWS le risorse necessarie per consentirti di connetterti ai tuoi cluster.
- **`AWS KMS`**autorizzazioni: vengono utilizzate durante le chiamate API per convalidare le risorse passate in una richiesta. Sono necessarie per consentire ad Amazon MSK di utilizzare la chiave passata con il cluster Amazon MSK.
- **`CloudWatch Logs, Amazon S3, and Amazon Data Firehose`**autorizzazioni: sono necessarie per consentire ad Amazon MSK di garantire che le destinazioni di consegna dei log siano raggiungibili e che siano valide per l'utilizzo dei log da parte dei broker.
- **`IAM`**autorizzazioni: sono necessarie per consentire ad Amazon MSK di creare un ruolo collegato al servizio nel tuo account e per consentirti di passare un ruolo di esecuzione del servizio ad Amazon MSK.

```
{
 "Version": "2012-10-17",
```

```
    "Statement": [{
      "Effect": "Allow",
      "Action": [
       "kafka:*",
       "ec2:DescribeSubnets",
       "ec2:DescribeVpcs",
       "ec2:DescribeSecurityGroups",
       "ec2:DescribeRouteTables",
       "ec2:DescribeVpcEndpoints",
       "ec2:DescribeVpcAttribute",
       "kms:DescribeKey",
       "kms:CreateGrant",
       "logs:CreateLogDelivery",
       "logs:GetLogDelivery",
       "logs:UpdateLogDelivery",
       "logs:DeleteLogDelivery",
       "logs:ListLogDeliveries",
       "logs:PutResourcePolicy",
       "logs:DescribeResourcePolicies",
       "logs:DescribeLogGroups",
       "S3:GetBucketPolicy",
       "firehose:TagDeliveryStream"
      ],
      "Resource": "*"
     },
     {
      "Effect": "Allow",
      "Action": [
       "ec2:CreateVpcEndpoint"
      ],
      "Resource": [
       "arn:*:ec2:*:*:vpc/*",
       "arn:*:ec2:*:*:subnet/*",
       "arn:*:ec2:*:*:security-group/*"
      ]
     },
     {
      "Effect": "Allow",
      "Action": [
       "ec2:CreateVpcEndpoint"
      ],
      "Resource": [
       "arn:*:ec2:*:*:vpc-endpoint/*"
      ],
```

```
        "Condition": {
         "StringEquals": {
          "aws:RequestTag/AWSMSKManaged": "true"
         },
         "StringLike": {
          "aws:RequestTag/ClusterArn": "*"
         }
        }
       },
       {
        "Effect": "Allow",
        "Action": [
         "ec2:CreateTags"
        ],
        "Resource": "arn:*:ec2:*:*:vpc-endpoint/*",
        "Condition": {
         "StringEquals": {
          "ec2:CreateAction": "CreateVpcEndpoint"
         }
        }
       },
       {
        "Effect": "Allow",
        "Action": [
         "ec2:DeleteVpcEndpoints"
        ],
        "Resource": "arn:*:ec2:*:*:vpc-endpoint/*",
        "Condition": {
         "StringEquals": {
          "ec2:ResourceTag/AWSMSKManaged": "true"
         },
         "StringLike": {
          "ec2:ResourceTag/ClusterArn": "*"
         }
        }
       },
       {
        "Effect": "Allow",
        "Action": "iam:PassRole",
        "Resource": "*",
        "Condition": {
         "StringEquals": {
          "iam:PassedToService": "kafka.amazonaws.com"
         }
```

```
       }
      },
      {
       "Effect": "Allow",
       "Action": "iam:CreateServiceLinkedRole",
       "Resource": "arn:aws:iam::*:role/aws-service-role/kafka.amazonaws.com/
AWSServiceRoleForKafka*",
       "Condition": {
        "StringLike": {
         "iam:AWSServiceName": "kafka.amazonaws.com"
        }
       }
      },
      {
       "Effect": "Allow",
       "Action": [
        "iam:AttachRolePolicy",
        "iam:PutRolePolicy"
       ],
       "Resource": "arn:aws:iam::*:role/aws-service-role/kafka.amazonaws.com/
AWSServiceRoleForKafka*"
      },
      {
       "Effect": "Allow",
       "Action": "iam:CreateServiceLinkedRole",
       "Resource": "arn:aws:iam::*:role/aws-service-role/delivery.logs.amazonaws.com/
AWSServiceRoleForLogDelivery*",
       "Condition": {
        "StringLike": {
         "iam:AWSServiceName": "delivery.logs.amazonaws.com"
        }
       }
      }

     ]
    }
```

AWS politica gestita: AmazonMSK Access ReadOnly

Questa policy concede autorizzazioni di sola lettura che consentono agli utenti di visualizzare informazioni in Amazon MSK. I principali ai quali è collegata questa policy non possono effettuare aggiornamenti o eliminare risorse esistenti, né possono creare nuove risorse Amazon MSK. Ad esempio, i principali con queste autorizzazioni possono visualizzare l'elenco dei cluster e delle

configurazioni associati al proprio account, ma non possono modificare la configurazione o le impostazioni di alcun cluster. Le autorizzazioni in questa policy sono raggruppate come segue:

- **`Amazon MSK`**autorizzazioni: consentono di elencare le risorse Amazon MSK, descriverle e ottenere informazioni su di esse.
- **`Amazon EC2`**autorizzazioni: vengono utilizzate per descrivere Amazon VPC, le sottoreti, i gruppi di sicurezza e gli ENI associati a un cluster.
- **`AWS KMS`**autorizzazione: viene utilizzata per descrivere la chiave associata al cluster.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "kafka:Describe*",
                "kafka:List*",
                "kafka:Get*",
                "ec2:DescribeNetworkInterfaces",
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeSubnets",
                "ec2:DescribeVpcs",
                "kms:DescribeKey"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

AWS politica gestita: KafkaServiceRolePolicy

Non puoi collegarti KafkaServiceRolePolicy alle tue entità IAM. Questa policy è collegata a un ruolo collegato ai servizi che consente ad Amazon MSK di eseguire operazioni come la gestione degli endpoint VPC (connettori) nei cluster MSK, la gestione delle interfacce di rete e la gestione delle credenziali del cluster con AWS Secrets Manager. Per ulteriori informazioni, consulta the section called “Ruoli collegati ai servizi”.

AWS politica gestita: AWSMSKReplicatorExecutionRole

La `AWSMSKReplicatorExecutionRole` policy concede le autorizzazioni al replicatore Amazon MSK per replicare i dati tra cluster MSK. Le autorizzazioni in questa policy sono raggruppate come segue:

- **`cluster`**— Concede ad Amazon MSK Replicator le autorizzazioni per connettersi al cluster utilizzando l'autenticazione IAM. Concede inoltre le autorizzazioni per descrivere e modificare il cluster.
- **`topic`**— Concede ad Amazon MSK Replicator le autorizzazioni per descrivere, creare e modificare un argomento e per modificare la configurazione dinamica dell'argomento.
- **`consumer group`**— Concede ad Amazon MSK Replicator le autorizzazioni per descrivere e modificare gruppi di consumatori, leggere e scrivere dati da un cluster MSK e eliminare argomenti interni creati dal replicatore.

```
{
 "Version": "2012-10-17",
 "Statement": [
  {
   "Sid": "ClusterPermissions",
   "Effect": "Allow",
   "Action": [
    "kafka-cluster:Connect",
    "kafka-cluster:DescribeCluster",
    "kafka-cluster:AlterCluster",
    "kafka-cluster:DescribeTopic",
    "kafka-cluster:CreateTopic",
    "kafka-cluster:AlterTopic",
    "kafka-cluster:WriteData",
    "kafka-cluster:ReadData",
    "kafka-cluster:AlterGroup",
    "kafka-cluster:DescribeGroup",
    "kafka-cluster:DescribeTopicDynamicConfiguration",
    "kafka-cluster:AlterTopicDynamicConfiguration",
    "kafka-cluster:WriteDataIdempotently"
   ],
   "Resource": [
    "arn:aws:kafka:*:*:cluster/*"
   ]
  },
```

```
  {
   "Sid": "TopicPermissions",
   "Effect": "Allow",
   "Action": [
    "kafka-cluster:DescribeTopic",
    "kafka-cluster:CreateTopic",
    "kafka-cluster:AlterTopic",
    "kafka-cluster:WriteData",
    "kafka-cluster:ReadData",
    "kafka-cluster:DescribeTopicDynamicConfiguration",
    "kafka-cluster:AlterTopicDynamicConfiguration",
    "kafka-cluster:AlterCluster"
   ],
   "Resource": [
    "arn:aws:kafka:*:*:topic/*/*"
   ]
  },
  {
   "Sid": "GroupPermissions",
   "Effect": "Allow",
   "Action": [
    "kafka-cluster:AlterGroup",
    "kafka-cluster:DescribeGroup"
   ],
   "Resource": [
    "arn:aws:kafka:*:*:group/*/*"
   ]
  }
 ]
}
```

Amazon MSK aggiorna le politiche AWS gestite

Visualizza i dettagli sugli aggiornamenti delle politiche AWS gestite per Amazon MSK da quando questo servizio ha iniziato a tracciare queste modifiche.

Modifica	Descrizione	Data
WriteDataIdempotently autorizzazione aggiunta a AWSMSKReplicatorEx	Amazon MSK ha aggiunto WriteDataIdempoten tly l'autorizzazione alla AWSMSKReplicatorEx	12 marzo 2024

Modifica	Descrizione	Data
ecutionRole: aggiornamento a una politica esistente	ecutionRole policy per supportare la replica dei dati tra cluster MSK.	
AWSMSKReplicatorEx ecutionRole: nuova policy	Amazon MSK ha aggiunto una AWSMSKReplicatorEx ecutionRole policy per supportare Amazon MSK Replicator.	4 dicembre 2023
AmazonMSK FullAccess: aggiornamento a una politica esistente	Amazon MSK ha aggiunto le autorizzazioni per supportare il replicatore Amazon MSK.	28 settembre 2023
KafkaServiceRolePolicy: aggiornamento a una policy esistente	Amazon MSK ha aggiunto le autorizzazioni per supportar e la connettività privata multi-VPC.	8 marzo 2023
AmazonMSK FullAccess: aggiornamento a una politica esistente	Amazon MSK ha aggiunto nuove autorizzazioni Amazon EC2 per consentire la connessione a un cluster.	30 novembre 2021
AmazonMSK FullAccess: aggiornamento a una politica esistente	Amazon MSK ha aggiunto una nuova autorizzazione per consentirgli di descrivere le tabelle di routing di Amazon EC2.	19 novembre 2021
Amazon MSK ha iniziato a tenere traccia delle modifiche	Amazon MSK ha iniziato a tracciare le modifiche per le sue politiche AWS gestite.	19 novembre 2021

Risoluzione dei problemi relativi all'identità e all'accesso di Amazon MSK

Utilizza le informazioni seguenti per eseguire la diagnosi e risolvere i problemi comuni che possono verificarsi durante l'utilizzo di Amazon MSK e IAM.

Argomenti

Non dispongo dell'autorizzazione per eseguire un'operazione in Amazon MSK

Se ti AWS Management Console dice che non sei autorizzato a eseguire un'azione, devi contattare l'amministratore per ricevere assistenza. L'amministratore è colui che ti ha fornito le credenziali di accesso.

L'errore di esempio seguente si verifica quando l'utente IAM `mateojackson` cerca di utilizzare la console per eliminare un cluster senza disporre delle autorizzazioni `kafka:DeleteCluster`.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 kafka:DeleteCluster on resource: purchaseQueriesCluster
```

In questo caso, Mateo richiede al suo amministratore di aggiornare le policy per poter accedere alla risorsa `purchaseQueriesCluster` utilizzando l'azione `kafka:DeleteCluster`.

Autenticazione e autorizzazione per le API di Apache Kafka

Puoi utilizzare IAM per autenticare i client e consentire o rifiutare le operazioni di Apache Kafka. In alternativa, è possibile utilizzare TLS oppure SASL/SCRAM per autenticare i client e le ACL Apache Kafka per consentire o rifiutare le operazioni.

Per informazioni su come controllare chi può eseguire le operazioni di Amazon MSK sul tuo cluster, consulta la pagina the section called "Autenticazione e autorizzazione per le API di Amazon MSK".

Argomenti

- ACL Apache Kafka

Controllo degli accessi IAM

Il Controllo degli accessi IAM per Amazon MSK ti consente di gestire sia l'autenticazione sia l'autorizzazione per il cluster MSK. Ciò elimina la necessità di utilizzare meccanismi separati per l'autenticazione e l'autorizzazione. Ad esempio, quando un client tenta di scrivere sul cluster, Amazon MSK utilizza IAM per verificare se tale client è un'identità autenticata e se è autorizzato a produrre nel cluster. Il controllo degli accessi IAM funziona per client Java e non Java, inclusi i client Kafka scritti in Python, Go e.NET. JavaScript

Amazon MSK registra gli eventi di accesso in modo da poterli controllare. Per ulteriori informazioni, consulta the section called “CloudTrail eventi”.

Per rendere possibile il Controllo degli accessi IAM, Amazon MSK apporta piccole modifiche al codice sorgente di Apache Kafka. Queste modifiche non causeranno differenze evidenti nella tua esperienza con Apache Kafka.

> **Important**
>
> Il controllo degli accessi IAM non si applica ai nodi Apache. ZooKeeper Per ulteriori informazioni sul controllo degli accessi a tali nodi, consulta la pagina the section called “Controllo dell'accesso ad Apache ZooKeeper”.

> **Important**
>
> L'impostazione `allow.everyone.if.no.acl.found` di Apache Kafka non ha effetto se il cluster utilizza il Controllo degli accessi IAM.

> **Important**
>
> Puoi richiamare le API delle ACL di Apache Kafka per un cluster MSK che utilizza il Controllo degli accessi IAM. Tuttavia, gli ACL di Apache Kafka non hanno alcun effetto sull'autorizzazione per i ruoli IAM. Per il controllo degli accessi per i ruoli IAM è necessario utilizzare le policy IAM.

Come funziona il Controllo degli accessi IAM per Amazon MSK

Per utilizzare il Controllo degli accessi IAM per Amazon MSK, esegui i seguenti passaggi, descritti nel dettaglio nel resto di questa sezione.

- the section called “Creazione un cluster che utilizza il Controllo degli accessi IAM”
- the section called “Configurazione dei client per il Controllo degli accessi IAM”
- the section called “Creazione di policy di autorizzazione”
- the section called “Recupero dei broker di bootstrap per il Controllo degli accessi IAM”

Creazione un cluster che utilizza il Controllo degli accessi IAM

Questa sezione spiega come utilizzare l' AWS Management Console, l'API o il AWS CLI per creare un cluster che utilizza il controllo degli accessi IAM. Per informazioni su come attivare il Controllo degli accessi IAM per un cluster esistente, consulta la pagina the section called “Aggiornamento della sicurezza”.

Utilizza il AWS Management Console per creare un cluster che utilizza il controllo degli accessi IAM

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli Create cluster (Crea cluster).
3. Scegli Crea cluster con impostazioni personalizzate.
4. Nella sezione Autenticazione, scegli Controllo degli accessi IAM.
5. Completa il resto del flusso di lavoro per creare un cluster.

Utilizza l'API o il AWS CLI per creare un cluster che utilizza il controllo degli accessi IAM

- Per creare un cluster con il controllo degli accessi IAM abilitato, utilizza l'CreateClusterAPI o il comando CLI create-cluster e passa il seguente JSON per il parametro:. `ClientAuthentication` `"ClientAuthentication": { "Sasl": { "Iam": { "Enabled": true } }`

Configurazione dei client per il Controllo degli accessi IAM

Per consentire ai client di comunicare con un cluster MSK che utilizza il controllo degli accessi IAM, è possibile utilizzare uno di questi meccanismi:

- Configurazione del client non Java utilizzando il meccanismo SASL_OAUTHBEARER
- Configurazione del client Java utilizzando il meccanismo SASL_OAUTHBEARER o AWS_MSK_IAM

Utilizzo del meccanismo SASL_OAUTHBEARER per configurare IAM

1. Modifica il tuo file di configurazione client.properties usando la sintassi evidenziata nel client Python Kafka di esempio riportato di seguito come guida. Le modifiche alla configurazione sono simili in altri linguaggi.

```
#!/usr/bin/python3from kafka import KafkaProducer
from kafka.errors import KafkaError
import socket
import time
from aws_msk_iam_sasl_signer import MSKAuthTokenProvider

class MSKTokenProvider():
    def token(self):
        token, _ = MSKAuthTokenProvider.generate_auth_token('<my aws region>')
        return token

tp = MSKTokenProvider()

producer = KafkaProducer(
    bootstrap_servers='<my bootstrap string>',
    security_protocol='SASL_SSL',
    sasl_mechanism='OAUTHBEARER',
    sasl_oauth_token_provider=tp,
    client_id=socket.gethostname(),
)

topic = "<my-topic>"
while True:
    try:
        inp=input(">")
        producer.send(topic, inp.encode())
        producer.flush()
        print("Produced!")
    except Exception:
        print("Failed to send message:", e)
```

```
producer.close()
```

2. Scarica la libreria di supporto per il linguaggio di configurazione scelto e segui le istruzioni nella sezione Nozioni di base sulla home page della libreria del linguaggio.
   - JavaScript: https://github.com/aws/aws-msk-iam-sasl-signer-js#getting-started
   - Python: https://github.com/aws/aws-msk-iam-sasl-signer-python#get-started
   - Go: https://github.com/aws/aws-msk-iam-sasl-signer-go#getting-started
   - .NET: https://github.com/aws/aws-msk-iam-sasl-signer-net#getting-started
   - JAVA: il supporto SASL_OAUTHBEARER per Java è disponibile tramite il file jar `aws-msk-iam-auth`

Utilizzo del meccanismo personalizzato AWS_MSK_IAM di MSK per configurare IAM

1. Aggiungi quanto segue al file `client.properties`. Sostituisci *<PATH_TO_TRUST_STORE_FILE>* con il percorso completo del file di truststore sul client.

   > **Note**
   >
   > Se non desideri utilizzare un certificato specifico, puoi rimuovere `ssl.truststore.location=`*<PATH_TO_TRUST_STORE_FILE>* dal tuo file `client.properties`. Se non specifichi un valore per `ssl.truststore.location`, il processo Java utilizza il certificato predefinito.

   ```
   ssl.truststore.location=<PATH_TO_TRUST_STORE_FILE>
   security.protocol=SASL_SSL
   sasl.mechanism=AWS_MSK_IAM
   sasl.jaas.config=software.amazon.msk.auth.iam.IAMLoginModule required;
   sasl.client.callback.handler.class=software.amazon.msk.auth.iam.IAMClientCallbackHandler
   ```

   Per utilizzare un profilo denominato creato per AWS le credenziali, includilo `awsProfileName="`*your profile name*`";` nel file di configurazione del client. Per informazioni sui profili denominati, consulta Profili denominati nella AWS CLI documentazione.

2. Scarica l'ultimo file JAR stabile aws-msk-iam-auth e inseriscilo nel percorso della classe. Se utilizzi Maven, aggiungi la seguente dipendenza, modificando il numero di versione secondo necessità:

```
<dependency>
    <groupId>software.amazon.msk</groupId>
    <artifactId>aws-msk-iam-auth</artifactId>
    <version>1.0.0</version>
</dependency>
```

Il plug-in client di Amazon MSK è open source con licenza Apache 2.0.

Creazione di policy di autorizzazione

Collega una policy di autorizzazione al ruolo IAM corrispondente al client. In una policy di autorizzazione, specifichi quali operazioni consentire o rifiutare per il ruolo. Se il tuo client è su un'istanza Amazon EC2, associa la policy di autorizzazione al ruolo IAM per quell'istanza Amazon EC2. In alternativa, puoi configurare il client per utilizzare un profilo denominato e quindi associare la policy di autorizzazione al ruolo per quel profilo denominato. the section called “Configurazione dei client per il Controllo degli accessi IAM” descrive come configurare un client per utilizzare un profilo denominato.

Per informazioni sulla creazione di una policy IAM, consulta la pagina Creating IAM policies.

Di seguito è riportato un esempio di politica di autorizzazione per un cluster denominato MyTestCluster. Per comprendere la semantica degli elementi `Action` e `Resource`, consulta la pagina the section called “Semantica delle operazioni e delle risorse”.

> **Important**
>
> Le modifiche apportate a una policy IAM si riflettono immediatamente nella AWS CLI e nelle API di IAM. Tuttavia, può trascorrere molto tempo prima che la modifica della policy abbia effetto. Nella maggior parte dei casi, le modifiche alle policy entrano in vigore in meno di un minuto. A volte le condizioni della rete possono aumentare il ritardo.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:Connect",
```

```
                "kafka-cluster:AlterCluster",
                "kafka-cluster:DescribeCluster"
            ],
            "Resource": [
                "arn:aws:kafka:us-east-1:0123456789012:cluster/MyTestCluster/
abcd1234-0123-abcd-5678-1234abcd-1"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:*Topic*",
                "kafka-cluster:WriteData",
                "kafka-cluster:ReadData"
            ],
            "Resource": [
                "arn:aws:kafka:us-east-1:0123456789012:topic/MyTestCluster/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:AlterGroup",
                "kafka-cluster:DescribeGroup"
            ],
            "Resource": [
                "arn:aws:kafka:us-east-1:0123456789012:group/MyTestCluster/*"
            ]
        }
    ]
}
```

Per informazioni su come creare una policy con elementi di operazione che corrispondano ai casi d'uso comuni di Apache Kafka, come la produzione e l'utilizzo di dati, consulta la pagina the section called “Casi di utilizzo comune”.

Per le versioni 2.8.0 e successive di Kafka, l'autorizzazione WriteDataIdempotently è obsoleta (KIP-679). Per impostazione predefinita, viene utilizzato `enable.idempotence = true`. Pertanto, per le versioni di Kafka 2.8.0 e successive, IAM non offre le stesse funzionalità delle ACL di Kafka. Non è possibile eseguire l'operazione `WriteDataIdempotently` su un argomento fornendo l'accesso `WriteData` solo a quell'argomento. Ciò non influisce sul caso in cui `WriteData` venga fornito a TUTTI gli argomenti. In tal caso, l'operazione `WriteDataIdempotently` è consentita.

Ciò è dovuto alle differenze nell'implementazione della logica IAM rispetto al modo in cui vengono implementate le ACL di Kafka.

Per ovviare a questo problema, consigliamo di utilizzare una policy simile all'esempio seguente:

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:Connect",
                "kafka-cluster:AlterCluster",
                "kafka-cluster:DescribeCluster",
                "kafka-cluster:WriteDataIdempotently"
            ],
            "Resource": [
                "arn:aws:kafka:us-east-1:0123456789012:cluster/MyTestCluster/
abcd1234-0123-abcd-5678-1234abcd-1"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:*Topic*",
                "kafka-cluster:WriteData",
                "kafka-cluster:ReadData"
            ],
            "Resource": [
                "arn:aws:kafka:us-east-1:0123456789012:topic/MyTestCluster/
abcd1234-0123-abcd-5678-1234abcd-1/TestTopic"
            ]
        }
    ]
}
```

In questo caso, `WriteData` consente le scritture sul `TestTopic`, mentre `WriteDataIdempotently` consente le scritture idempotenti sul cluster. È importante notare che `WriteDataIdempotently` è un'autorizzazione a livello di cluster. Non può essere utilizzata a livello di argomento. Se `WriteDataIdempotently` è limitato al livello di argomento, questa policy non funzionerà.

Recupero dei broker di bootstrap per il Controllo degli accessi IAM

Per informazioni, consulta the section called “Recupero dei broker di bootstrap”.

Semantica delle operazioni e delle risorse

In questa sezione viene illustrata la semantica degli elementi di operazione e risorsa che è possibile utilizzare in una policy di autorizzazione IAM. Per un esempio di policy, consulta the section called “Creazione di policy di autorizzazione”.

Azioni

La tabella seguente elenca le operazioni che è possibile includere in una policy di autorizzazione quando si utilizza il Controllo degli accessi IAM per Amazon MSK. Quando includi nella tua policy di autorizzazione un'operazione dalla colonna Operazione della tabella, devi includere anche le operazioni corrispondenti dalla colonna Operazioni richieste.

Azione	Descrizione	Operazioni necessarie	Risorse obbligatorie	Applicabile ai cluster serverless
`kafka-cluster:Connect`	Concede l'autorizzazione per connettersi e autenticarsi al cluster.	Nessuno	cluster	Sì
`kafka-cluster:DescribeCluster`	Concede l'autorizzazione per descrivere vari aspetti del cluster, equivalente all'ACL DESCRIBE CLUSTER di Apache Kafka.	`kafka-cluster:Connect`	cluster	Sì

Azione	Descrizione	Operazioni necessarie	Risorse obbligatorie	Applicabile ai cluster serverless
`kafka-cluster:AlterCluster`	Concede l'autorizzazione per modificare vari aspetti del cluster, equivalente all'ACL ALTER CLUSTER di Apache Kafka.	`kafka-cluster:Connect` `kafka-cluster:DescribeCluster`	cluster	No
`kafka-cluster:DescribeClusterDynamicConfiguration`	Concede l'autorizzazione per descrivere la configurazione dinamica di un cluster, equivalente all'ACL DESCRIBE_CONFIGS CLUSTER di Apache Kafka.	`kafka-cluster:Connect`	cluster	No

Azione	Descrizione	Operazioni necessarie	Risorse obbligatorie	Applicabile ai cluster serverless

`kafka-cluster:AlterClusterDynamicConfiguration`	Concede l'autorizzazione per modificare la configurazione dinamica di un cluster, equivalente all'ACL ALTER_CONFIGS CLUSTER di Apache Kafka.	`kafka-cluster:Connect`
		
		`kafka-cluster:DescribeClusterDynamicConfiguration`
			cluster	No

`kafka-cluster:WriteDataIdempotently`	Concede l'autorizzazione per scrivere dati in modo idempotente su un cluster, equivalente all'ACL IDEMPOTENT_WRITE CLUSTER di Apache Kafka.	`kafka-cluster:Connect`
		
		`kafka-cluster:WriteData`
			cluster	Sì

`kafka-cluster:CreateTopic`	Concede l'autorizzazione per creare argomenti in un cluster, equivalente all'ACL CREATE CLUSTER/TOPIC di Apache Kafka.	`kafka-cluster:Connect`	topic	Sì

Azione	Descrizione	Operazioni necessarie	Risorse obbligatorie	Applicabile ai cluster serverless
`kafka-cluster:DescribeTopic`	Concede l'autorizzazione per descrivere gli argomenti in un cluster, equivalente all'ACL DESCRIBE TOPIC di Apache Kafka.	`kafka-cluster:Connect`	topic	Sì
`kafka-cluster:AlterTopic`	Concede l'autorizzazione per modificare gli argomenti in un cluster, equivalente all'ACL ALTER TOPIC di Apache Kafka.	`kafka-cluster:Connect` `kafka-cluster:DescribeTopic`	topic	Sì
`kafka-cluster:DeleteTopic`	Concede l'autorizzazione per eliminare gli argomenti in un cluster, equivalente all'ACL DELETE TOPIC di Apache Kafka.	`kafka-cluster:Connect` `kafka-cluster:DescribeTopic`	topic	Sì

Azione	Descrizione	Operazioni necessarie	Risorse obbligatorie	Applicabile ai cluster serverless
`kafka-cluster:DescribeTopicDynamicConfiguration`	Concede l'autorizzazione per descrivere la configurazione dinamica degli argomenti in un cluster, equivalente all'ACL DESCRIBE_CONFIGS TOPIC di Apache Kafka.	`kafka-cluster:Connect`	topic	Sì

`kafka-cluster:AlterTopicDynamicConfiguration`	Concede l'autorizzazione per modificare la configurazione dinamica degli argomenti in un cluster, equivalente all'ACL ALTER_CONFIGS TOPIC di Apache Kafka.	`kafka-cluster:Connect`
		
		`kafka-cluster:DescribeTopicDynamicConfiguration`
			topic	Sì

Azione	Descrizione	Operazioni necessarie	Risorse obbligatorie	Applicabile ai cluster serverless

`kafka-cluster:ReadData`	Concede l'autorizzazione per leggere i dati da argomenti in un cluster, equivalente all'ACL READ TOPIC di Apache Kafka.	`kafka-cluster:Connect`
		`kafka-cluster:DescribeTopic`
		`kafka-cluster:AlterGroup`
			topic	Sì

`kafka-cluster:WriteData`	Concede l'autorizzazione per scrivere dati su argomenti in un cluster, equivalente all'ACL WRITE TOPIC di Apache Kafka	`kafka-cluster:Connect` `kafka-cluster:DescribeTopic`	topic	Sì
`kafka-cluster:DescribeGroup`	Concede l'autorizzazione per descrivere i gruppi in un cluster, equivalente all'ACL DESCRIBE GROUP di Apache Kafka.	`kafka-cluster:Connect`	gruppo	Sì

Azione	Descrizione	Operazioni necessarie	Risorse obbligatorie	Applicabile ai cluster serverless
`kafka-cluster:AlterGroup`	Concede l'autorizzazione per unire dei gruppi all'interno di un cluster, equivalente all'ACL READ GROUP di Apache Kafka.	`kafka-cluster:Connect` `kafka-cluster:DescribeGroup`	gruppo	Sì
`kafka-cluster:DeleteGroup`	Concede l'autorizzazione per eliminare gruppi all'interno di un cluster, equivalente all'ACL DELETE GROUP di Apache Kafka.	`kafka-cluster:Connect` `kafka-cluster:DescribeGroup`	gruppo	Sì
`kafka-cluster:DescribeTransactionalId`	Concede l'autorizzazione per descrivere gli ID transazionali in un cluster, equivalente all'ACL DESCRIBE TRANSACTIONAL_ID di Apache Kafka.	`kafka-cluster:Connect`	transactional-id	Sì

Azione	Descrizione	Operazioni necessarie	Risorse obbligatorie	Applicabile ai cluster serverless

`kafka-cluster:AlterTransactionalId`	Concede l'autorizzazione per modificare gli ID transazionali in un cluster, equivalente all'ACL WRITE TRANSACTIONAL_ID di Apache Kafka.	`kafka-cluster:Connect`
		`kafka-cluster:DescribeTransactionalId`
		`kafka-cluster:WriteData`
			transactional-id	Sì

In un'operazione, dopo i due punti, è possibile utilizzare qualsiasi quantità di caratteri jolly asterisco (*). Di seguito vengono mostrati gli esempi.

- `kafka-cluster:*Topic` sta per `kafka-cluster:CreateTopic`, `kafka-cluster:DescribeTopic`, `kafka-cluster:AlterTopic` e `kafka-cluster:DeleteTopic`. Non include `kafka-cluster:DescribeTopicDynamicConfiguration` o `kafka-cluster:AlterTopicDynamicConfiguration`.
- `kafka-cluster:*` indica tutte le autorizzazioni.

Risorse

La tabella seguente mostra i quattro tipi di risorse che è possibile utilizzare in una policy di autorizzazione quando si utilizza il Controllo degli accessi IAM per Amazon MSK. Puoi ottenere il cluster Amazon Resource Name (ARN) da o utilizzando l'DescribeClusterAPI AWS Management Console o il comando AWS CLI describe-cluster. È quindi possibile utilizzare l'ARN del cluster per creare ARN per argomenti, gruppi e ID transazionale. Per specificare una risorsa nella policy di autorizzazione, utilizza l'ARN della risorsa.

Risorsa	Formato ARN
Cluster	arn:aws:kafka:*region*:*account-id* :cluster/*cluster-name* /*cluster-uuid*
Argomento	arn:aws:kafka:*region*:*account-id* :topic/*cluster-name* /*cluster-uuid* /*topic-name*
Group (Gruppo)	arn:aws:kafka:*region*:*account-id* :group/*cluster-name* /*cluster-uuid* /*group-name*
ID transazionale	arn:aws:kafka:*region*:*account-id* :transactional-id/*cluster-name* /*cluster-uuid* /*transactional-id*

È possibile utilizzare qualsiasi quantità di caratteri jolly asterisco (*) in qualsiasi punto dell'ARN successivo a `:cluster/`, `:topic/`, `:group/` e `:transactional-id/`. Di seguito sono riportati alcuni esempi di come utilizzare il carattere jolly asterisco (*) per fare riferimento a più risorse:

- `arn:aws:kafka:us-east-1:0123456789012:topic/MyTestCluster/*`: tutti gli argomenti di qualsiasi cluster denominato MyTestCluster, indipendentemente dall'UUID del cluster.
- `arn:aws:kafka:us-east-1:0123456789012:topic/MyTestCluster/abcd1234-0123-abcd-5678-1234abcd-1/*_test`: tutti gli argomenti il cui nome termina con «_test» nel cluster il cui nome è MyTestCluster e il cui UUID è abcd1234-0123-abcd-5678-1234abcd-1.
- `arn:aws:kafka:us-east-1:0123456789012:transactional-id/MyTestCluster/*/5555abcd-1111-abcd-1234-abcd1234-1`: tutte le transazioni il cui ID transazionale è MyTestCluster 5555abcd-1111-abcd-1234-abcd1234-1, in tutte le incarnazioni di un cluster denominato nel tuo account. Ciò significa che se si crea un cluster denominato MyTestCluster, quindi lo si elimina e quindi si crea un altro cluster con lo stesso nome, è possibile utilizzare questa risorsa ARN per rappresentare lo stesso ID delle transazioni su entrambi i cluster. Tuttavia, il cluster eliminato non è accessibile.

Casi di utilizzo comune

La prima colonna della tabella seguente mostra alcuni casi d'uso comuni. Per autorizzare un client a eseguire un determinato caso d'uso, includi le operazioni richieste per tale caso d'uso nella policy di autorizzazione del client e imposta `Effect` su `Allow`.

Per informazioni su tutte le operazioni che fanno parte del Controllo degli accessi IAM per Amazon MSK, consulta la pagina the section called “Semantica delle operazioni e delle risorse”.

> **Note**
>
> Le operazioni non sono consentite per impostazione predefinita. È necessario consentire esplicitamente ogni operazione che si desidera autorizzare il client a eseguire.

Caso d'uso	Operazioni necessarie
Admin	`kafka-cluster:*`
Creazione di un argomento	`kafka-cluster:Connect` `kafka-cluster:CreateTopic`

Produzione di dati	`kafka-cluster:Connect`
	`kafka-cluster:DescribeTopic`
	`kafka-cluster:WriteData`

Utilizzo di dati	`kafka-cluster:Connect`
	`kafka-cluster:DescribeTopic`
	`kafka-cluster:DescribeGroup`
	`kafka-cluster:AlterGroup`
	`kafka-cluster:ReadData`

Produzione di dati in modo idempotente	`kafka-cluster:Connect`
	`kafka-cluster:DescribeTopic`
	`kafka-cluster:WriteData`
	`kafka-cluster:WriteDataIdempotently`

Caso d'uso	Operazioni necessarie

Produzione di dati in modo transazionale	`kafka-cluster:Connect`
	`kafka-cluster:DescribeTopic`
	`kafka-cluster:WriteData`
	`kafka-cluster:DescribeTransactionalId`
	`kafka-cluster:AlterTransactionalId`

Descrizione della configurazione di un cluster	`kafka-cluster:Connect` `kafka-cluster:DescribeClusterDynamicConfiguration`

Aggiornamento della configurazione di un cluster	`kafka-cluster:Connect`
	`kafka-cluster:DescribeClusterDynamicConfiguration`
	`kafka-cluster:AlterClusterDynamicConfiguration`

Descrizione della configurazione di un argomento	`kafka-cluster:Connect` `kafka-cluster:DescribeTopicDynamicConfiguration`

Aggiornamento della configurazione di un argomento	`kafka-cluster:Connect`
	`kafka-cluster:DescribeTopicDynamicConfiguration`
	`kafka-cluster:AlterTopicDynamicConfiguration`

Caso d'uso	Operazioni necessarie

Modifica di un argomento	`kafka-cluster:Connect`
	`kafka-cluster:DescribeTopic`
	`kafka-cluster:AlterTopic`

Autenticazione TLS reciproca

Puoi abilitare l'autenticazione client con TLS per le connessioni dalle tue applicazioni ai broker Amazon MSK. Per utilizzare l'autenticazione client, è necessario un CA privata AWS. CA privata AWS Possono appartenere allo Account AWS stesso cluster o a un account diverso. Per informazioni su CA privata AWS s, vedere Creazione e gestione di un CA privata AWS.

> **Note**
>
> L'autenticazione TLS non è disponibile nelle regioni di Pechino e Ningxia.

Amazon MSK non supporta le liste di revoche di certificati (CRL). Per controllare l'accesso agli argomenti del cluster o bloccare i certificati compromessi, utilizza gli ACL e i gruppi di sicurezza di Apache Kafka. AWS Per ulteriori informazioni sull'utilizzo delle ACL di Apache Kafka, consulta la pagina the section called “ACL Apache Kafka”.

Questo argomento contiene le sezioni seguenti:

- Per creare un cluster che supporta l'autenticazione client
- Per impostare un client per utilizzare l'autenticazione
- Per produrre e consumare messaggi utilizzando l'autenticazione

Per creare un cluster che supporta l'autenticazione client

Questa procedura mostra come abilitare l'autenticazione del client utilizzando un. CA privata AWS

> **Note**
>
> Si consiglia vivamente di utilizzare Independent CA privata AWS per ogni cluster MSK quando si utilizza il TLS reciproco per controllare l'accesso. In questo modo si assicurerà che i certificati TLS firmati dalle PCA si autentichino solo con un singolo cluster MSK.

1. Crea un file denominato `clientauthinfo.json` con i seguenti contenuti. Sostituire *`Private-CA-ARN`* con l'ARN del PCA.

```
{
   "Tls": {
       "CertificateAuthorityArnList": ["Private-CA-ARN"]
    }
}
```

2. Crea un file denominato `brokernodegroupinfo.json` come descritto in the section called “Creazione di un cluster utilizzando AWS CLI”.
3. L'autenticazione client richiede di abilitare anche la crittografia dei dati in transito tra client e broker. Crea un file denominato `encryptioninfo.json` con i seguenti contenuti. Sostituisci *`KMS-Key-Arn`* con l'ARN della chiave KMS. Puoi impostare `ClientBroker` su `TLS` o `TLS_PLAINTEXT`.

```
{
   "EncryptionAtRest": {
       "DataVolumeKMSKeyId": "KMS-Key-ARN"
    },
   "EncryptionInTransit": {
        "InCluster": true,
        "ClientBroker": "TLS"
    }
}
```

   Per ulteriori informazioni sulla crittografia, consulta the section called “Crittografia”.
4. Su una macchina su cui è AWS CLI installato, esegui il comando seguente per creare un cluster con l'autenticazione e la crittografia in transito abilitate. Salva l'ARN del cluster fornito nella risposta.

```
aws kafka create-cluster --cluster-name "AuthenticationTest" --broker-node-group-
info file://brokernodegroupinfo.json --encryption-info file://encryptioninfo.json
 --client-authentication file://clientauthinfo.json --kafka-version "{YOUR KAFKA
 VERSION}" --number-of-broker-nodes 3
```

Per impostare un client per utilizzare l'autenticazione

1. Crea un'istanza Amazon EC2 da utilizzare come un computer client. Per semplicità, creare questa istanza nello stesso VPC utilizzato per il cluster. Consulta the section called “Passaggio 3: creazione di un computer client” per un esempio di come creare un computer client di questo tipo.
2. Creazione di un argomento. Per un esempio, consulta le istruzioni in the section called “Passaggio 4: creazione di un argomento”.
3. Su un computer in cui è AWS CLI installato, esegui il comando seguente per ottenere i broker bootstrap del cluster. Sostituire *Cluster-ARN* con l'ARN del cluster.

   ```
   aws kafka get-bootstrap-brokers --cluster-arn Cluster-ARN
   ```

   Salvare la stringa associata a `BootstrapBrokerStringTls` nella risposta.
4. Sul computer client, eseguire il comando seguente per utilizzare il truststore JVM per creare il truststore client. Se il percorso JVM è diverso, modificare il comando di conseguenza.

   ```
   cp /usr/lib/jvm/java-1.8.0-openjdk-1.8.0.201.b09-0.amzn2.x86_64/jre/lib/security/
   cacerts kafka.client.truststore.jks
   ```
5. Sul computer client, eseguire il comando seguente per creare una chiave privata per il client. Sostituire *Distinguished-Name*, *Example-Alias*, *Your-Store-Pass* e *Your-Key-Pass* con stringhe prescelte.

   ```
   keytool -genkey -keystore kafka.client.keystore.jks -validity 300 -storepass Your-
   Store-Pass -keypass Your-Key-Pass -dname "CN=Distinguished-Name" -alias Example-
   Alias -storetype pkcs12
   ```
6. Sul computer client, eseguire il comando seguente per creare una richiesta di certificato con la chiave privata creata nella fase precedente.

```
keytool -keystore kafka.client.keystore.jks -certreq -file client-cert-sign-request
 -alias Example-Alias -storepass Your-Store-Pass -keypass Your-Key-Pass
```

7. Aprire il file `client-cert-sign-request` e accertarsi che inizi con `-----BEGIN CERTIFICATE REQUEST-----` e termini con `-----END CERTIFICATE REQUEST-----`. Se inizia con `-----BEGIN NEW CERTIFICATE REQUEST-----`, eliminare la parola `NEW` (e il singolo spazio che la segue) dall'inizio e dalla fine del file.
8. Su un computer su cui è AWS CLI installato, esegui il comando seguente per firmare la richiesta di certificato. Sostituire *`Private-CA-ARN`* con l'ARN del PCA. Se lo si desidera, è possibile modificare il valore di validità. In questo esempio viene utilizzato 300.

```
aws acm-pca issue-certificate --certificate-authority-arn Private-CA-ARN --csr
 fileb://client-cert-sign-request --signing-algorithm "SHA256WITHRSA" --validity
 Value=300,Type="DAYS"
```

Salvare il certificato ARN fornito nella risposta.

> **Note**
>
> Per recuperare il certificato client, utilizza il comando `acm-pca get-certificate` e specifica l'ARN del certificato. Per ulteriori informazioni, consulta la sezione get-certificate nella documentazione di riferimento alla AWS CLI .

9. Esegui il comando seguente per ottenere il certificato CA privata AWS firmato per te. Sostituire *`Certificate-ARN`* con l'ARN ottenuto dalla risposta al comando precedente.

```
aws acm-pca get-certificate --certificate-authority-arn Private-CA-ARN --
certificate-arn Certificate-ARN
```

10. Dal risultato JSON dell'esecuzione del comando precedente, copiare le stringhe associate a `Certificate` e `CertificateChain`. Incolla queste due stringhe in un nuovo file denominato signed-certificate-from-acm. Incollare innanzitutto la stringa associata a `Certificate`, seguita dalla stringa associata a `CertificateChain`. Sostituire i caratteri `\n` con nuove righe. Di seguito è riportata la struttura del file dopo aver incollato al suo interno il certificato e la catena di certificati.

```
-----BEGIN CERTIFICATE-----
...
```

```
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
...
-----END CERTIFICATE-----
```

11. Eseguire il comando seguente sul computer client per aggiungere questo certificato al keystore in modo da poterlo presentare quando si parla con i broker MSK.

```
keytool -keystore kafka.client.keystore.jks -import -file signed-certificate-from-
acm -alias Example-Alias -storepass Your-Store-Pass -keypass Your-Key-Pass
```

12. Crea un file denominato `client.properties` con i seguenti contenuti. Regolare le posizioni del truststore e del keystore sui percorsi in cui è stato salvato `kafka.client.truststore.jks`. Sostituisci i segnaposto *{YOUR KAFKA VERSION}* con la versione del tuo client Kafka.

```
security.protocol=SSL
ssl.truststore.location=/tmp/kafka_2.12-{YOUR KAFKA VERSION}/
kafka.client.truststore.jks
ssl.keystore.location=/tmp/kafka_2.12-{YOUR KAFKA VERSION}/
kafka.client.keystore.jks
ssl.keystore.password=Your-Store-Pass
ssl.key.password=Your-Key-Pass
```

Per produrre e consumare messaggi utilizzando l'autenticazione

1. Eseguire il comando seguente per creare un argomento. Il file denominato `client.properties` è quello creato nella procedura precedente.

```
<path-to-your-kafka-installation>/bin/kafka-topics.sh --create --bootstrap-
server BootstrapBroker-String --replication-factor 3 --partitions 1 --topic
 ExampleTopic --command-config client.properties
```

2. Eseguire il comando seguente per avviare un produttore della console. Il file denominato `client.properties` è quello creato nella procedura precedente.

```
<path-to-your-kafka-installation>/bin/kafka-console-producer.sh --bootstrap-
server BootstrapBroker-String --topic ExampleTopic --producer.config
 client.properties
```

3. In una nuova finestra di comando sul computer client, eseguire il comando seguente per avviare un consumatore della console.

```
<path-to-your-kafka-installation>/bin/kafka-console-consumer.sh --bootstrap-
server BootstrapBroker-String --topic ExampleTopic --consumer.config
 client.properties
```

4. Digitare i messaggi nella finestra del produttore e guardarli apparire nella finestra del consumatore.

Autenticazione delle credenziali di accesso con Secrets Manager AWS

Puoi controllare l'accesso ai tuoi cluster Amazon MSK utilizzando credenziali di accesso archiviate e protette tramite Secrets Manager. AWS L'archiviazione delle credenziali utente in Secrets Manager riduce il sovraccarico dell'autenticazione del cluster, ad esempio il controllo, l'aggiornamento e la rotazione delle credenziali. Secrets Manager consente inoltre di condividere le credenziali utente tra i cluster.

Questo argomento contiene le sezioni seguenti:

- Come funziona
- Configurazione dell'autenticazione SASL/SCRAM per un cluster Amazon MSK
- Operazioni con gli utenti
- Limitazioni

Come funziona

L'autenticazione delle credenziali di accesso di Amazon MSK supporta l'autenticazione SASL/SCRAM (Simple Authentication and Security Layer/Salted Challenge Response Authentication Mechanism). Per configurare l'autenticazione delle credenziali di accesso per un cluster, crea una risorsa segreta in AWS Secrets Manager e associa le credenziali di accesso a quel segreto.

SASL/SCRAM è definito in RFC 5802. SCRAM utilizza algoritmi di hashing protetti e non trasmette credenziali di accesso non crittografate tra client e server.

> **Note**
>
> Quando configuri l'autenticazione SASL/SCRAM per il cluster, Amazon MSK attiva la crittografia TLS per tutto il traffico tra client e broker.

Configurazione dell'autenticazione SASL/SCRAM per un cluster Amazon MSK

Per impostare un segreto in AWS Secrets Manager, segui il tutorial Creazione e recupero di un segreto nella Guida per l'utente di AWS Secrets Manager.

Tieni presente i seguenti requisiti quando crei un segreto per un cluster Amazon MSK:

- Per il tipo di segreto, scegli Altro tipo di segreto (es. chiave API).
- Il nome del segreto deve iniziare con il prefisso AmazonMSK_.
- È necessario utilizzare una AWS KMS chiave personalizzata esistente o creare una nuova AWS KMS chiave personalizzata per il segreto. Secrets Manager utilizza la AWS KMS chiave predefinita per un segreto per impostazione predefinita.

  > **Important**
  >
  > Un segreto creato con la AWS KMS chiave predefinita non può essere utilizzato con un cluster Amazon MSK.

- I dati delle credenziali di accesso devono essere nel seguente formato per inserire coppie chiave-valore utilizzando l'opzione Non crittografato.

  ```
  {
    "username": "alice",
    "password": "alice-secret"
  }
  ```

- Prendi nota del valore del nome della risorsa Amazon (ARN) del segreto.
-

  > **Important**
  >
  > Non è possibile associare un segreto di Secrets Manager a un cluster che supera i limiti descritti in the section called “ Dimensionamento corretto del cluster: numero di partizioni per broker”.

- Se si utilizza il AWS CLI per creare il segreto, specificare un ID chiave o un ARN per il `kms-key-id` parametro. Non specificare un alias.
- Per associare il segreto al cluster, utilizza la console Amazon MSK o l' BatchAssociateScramSecretoperazione.

> **Important**
>
> Quando associ un segreto a un cluster, Amazon MSK collega al segreto una policy delle risorse che consente al cluster di accedere e leggere i valori del segreto che hai definito. Questa policy delle risorse non dovrebbe essere modificata. In questo modo, è possibile impedire al cluster di accedere al segreto.

L'esempio seguente di input JSON per l'operazione `BatchAssociateScramSecret` associa un segreto a un cluster:

```
{
  "clusterArn" : "arn:aws:kafka:us-west-2:0123456789019:cluster/SalesCluster/
abcd1234-abcd-cafe-abab-9876543210ab-4",
  "secretArnList": [
    "arn:aws:secretsmanager:us-west-2:0123456789019:secret:AmazonMSK_MyClusterSecret"
  ]
}
```

Connessione al cluster con credenziali di accesso

Dopo aver creato un segreto e averlo collegato al cluster, è possibile collegare il client al cluster. I seguenti passaggi di esempio mostrano come connettere un client a un cluster che utilizza l'autenticazione SASL/SCRAM e come produrre e utilizzare un argomento di esempio.

1. *Esegui il comando seguente su un computer su cui è installata la AWS CLI, sostituendo clusterARN con l'ARN del tuo cluster.*

   ```
   aws kafka get-bootstrap-brokers --cluster-arn clusterARN
   ```

2. Per creare un argomento di esempio, esegui il comando seguente, sostituendo *BootstrapServerString* con uno degli endpoint del broker ottenuti nel passaggio precedente.

```
<path-to-your-kafka-installation>/bin/kafka-topics.sh --create --bootstrap-
server BootstrapServerString --replication-factor 3 --partitions 1 --topic
 ExampleTopicName
```

3. Sul tuo computer client, crea un file di configurazione JAAS che contenga le credenziali utente archiviate nel tuo segreto. Ad esempio, per l'utente alice, crea un file chiamato `users_jaas.conf` con il seguente contenuto.

```
KafkaClient {
   org.apache.kafka.common.security.scram.ScramLoginModule required
   username="alice"
   password="alice-secret";
};
```

4. Utilizza il seguente comando per esportare il file di configurazione JAAS come parametro di ambiente `KAFKA_OPTS`.

```
export KAFKA_OPTS=-Djava.security.auth.login.config=<path-to-jaas-file>/
users_jaas.conf
```

5. Nella directory `./tmp`, crea un file denominato `kafka.client.truststore.jks`.
6. Utilizza il comando seguente per copiare il file dell'archivio chiavi JDK dalla cartella `cacerts` JVM nel file `kafka.client.truststore.jks` creato nel passaggio precedente. Sostituisci *`JDKFolder`* con il nome della cartella JDK sull'istanza. Ad esempio, la tua cartella JDK potrebbe avere il nome `java-1.8.0-openjdk-1.8.0.201.b09-0.amzn2.x86_64`.

```
cp /usr/lib/jvm/JDKFolder/jre/lib/security/cacerts /tmp/kafka.client.truststore.jks
```

7. Nella directory `bin` di installazione di Apache Kafka, crea un file delle proprietà del client chiamato `client_sasl.properties` con il seguente contenuto. Questo file definisce il meccanismo e il protocollo SASL.

```
security.protocol=SASL_SSL
sasl.mechanism=SCRAM-SHA-512
ssl.truststore.location=<path-to-keystore-file>/kafka.client.truststore.jks
```

8. Recupera la stringa dei broker di bootstrap con il comando seguente. Sostituisci *`ClusterArn`*con l'Amazon Resource Name (ARN) del tuo cluster:

```
aws kafka get-bootstrap-brokers --cluster-arn ClusterArn
```

Dal risultato JSON del comando, salva il valore associato alla stringa denominata `BootstrapBrokerStringSaslScram`.

9. Per produrre l'argomento di esempio che hai creato, esegui il comando seguente sul computer client. Sostituisci *BootstrapBrokerStringSaslScram* con il valore recuperato nel passaggio precedente.

```
<path-to-your-kafka-installation>/bin/kafka-console-producer.sh --broker-
list BootstrapBrokerStringSaslScram --topic ExampleTopicName --producer.config
 client_sasl.properties
```

10. Per utilizzare l'argomento che hai creato, esegui il comando seguente sul tuo computer client. Sostituisci *BootstrapBrokerStringSaslScram* con il valore che hai ottenuto in precedenza.

```
<path-to-your-kafka-installation>/bin/kafka-console-consumer.sh --bootstrap-
server BootstrapBrokerStringSaslScram --topic ExampleTopicName --from-beginning --
consumer.config client_sasl.properties
```

Operazioni con gli utenti

Creazione di utenti: crea utenti nel tuo segreto come coppie chiave-valore. Quando si utilizza l'opzione Non crittografato nella console Secrets Manager, è necessario specificare i dati delle credenziali di accesso nel formato seguente.

```
{
  "username": "alice",
  "password": "alice-secret"
}
```

Revoca dell'accesso utente: per revocare le credenziali di accesso a un cluster di un utente, si consiglia di rimuovere o applicare un'ACL al cluster e successivamente annullare l'associazione del segreto. Ciò può essere dovuto ai motivi seguenti:

- La rimozione di un utente non chiude le connessioni esistenti.
- La propagazione delle modifiche al segreto richiede fino a 10 minuti.

Per ulteriori informazioni sull'utilizzo delle ACL con Amazon MSK, consulta la pagina ACL Apache Kafka.

Per i cluster che utilizzano ZooKeeper la modalità, si consiglia di limitare l'accesso ai ZooKeeper nodi per impedire agli utenti di modificare gli ACL. Per ulteriori informazioni, consulta Controllo dell'accesso ad Apache ZooKeeper.

Limitazioni

Quando utilizzi i segreti SCRAM, tieni presente le limitazioni seguenti:

- Amazon MSK supporta solo l'autenticazione SCRAM-SHA-512.
- Un cluster Amazon MSK può avere fino a 1.000 utenti.
- Devi usare an AWS KMS key con il tuo Secret. Non è possibile utilizzare un segreto che utilizza la chiave di crittografia Secrets Manager predefinita con Amazon MSK. Per ulteriori informazioni sulla creazione di una chiave KMS, consulta la pagina Creating symmetric encryption KMS keys.
- Non è possibile utilizzare una chiave KMS asimmetrica con Secrets Manager.
- È possibile associare fino a 10 segreti a un cluster alla volta utilizzando l' BatchAssociateScramSecretoperazione.
- Il nome dei segreti associati a un cluster Amazon MSK deve avere il prefisso AmazonMSK_.
- I segreti associati a un cluster Amazon MSK devono trovarsi nello stesso account e nella stessa AWS regione Amazon Web Services del cluster.

ACL Apache Kafka

Apache Kafka dispone di un autorizzatore collegabile e viene fornito con un'implementazione di autorizzazione. out-of-box Amazon MSK abilita questo provider di autorizzazioni nel file `server.properties` sui broker.

Gli ACL di Apache Kafka hanno il formato «Principal P è [Consentita/Negata] Operazione O dall'host H su qualsiasi risorsa R corrispondente a RP». ResourcePattern Se RP non corrisponde a una risorsa R specifica, R non dispone di ACL associati e pertanto nessuno, a parte i superuser, è autorizzato ad accedere a R. Per modificare questo comportamento di Apache Kafka, impostare la proprietà `allow.everyone.if.no.acl.found` su true. In Amazon MSK è impostata su true per impostazione predefinita. Ciò significa che nei cluster Amazon MSK, se non si impostano esplicitamente gli ACL su una risorsa, tutti i principali possono accedere a questa risorsa. Se si

abilitano gli ACL su una risorsa, l'accesso è consentito solo ai principali autorizzati. Se si desidera limitare l'accesso a un argomento e autorizzare un client utilizzando l'autenticazione reciproca TLS, aggiungere ACL utilizzando l'interfaccia della riga di comando del provider di autorizzazioni Apache Kafka. Per ulteriori informazioni sull'aggiunta, la rimozione e l'elenco di ACL, consulta Kafka Authorization Command Line Interface.

Oltre al client, è inoltre necessario concedere a tutti i broker l'accesso agli argomenti in modo che possano replicare i messaggi dalla partizione primaria. Se i broker non dispongono dell'accesso a un argomento, la replica per l'argomento non va a buon fine.

Per aggiungere o rimuovere l'accesso in lettura e scrittura a un argomento

1. Aggiungere i broker alla tabella ACL per consentire loro di leggere da tutti gli argomenti contenenti ACL. Per concedere ai broker l'accesso in lettura a un argomento, esegui il comando seguente su un computer client in grado di comunicare con il cluster MSK.

   Sostituire *Distinguished-Name* con il DNS di un broker bootstrap del cluster qualsiasi, quindi sostituire la stringa prima del primo periodo in questo nome distinto mediante un asterisco (*). Ad esempio, se il DNS di uno dei broker bootstrap del cluster è `b-6.mytestcluster.67281x.c4.kafka.us-east-1.amazonaws.com`, sostituire *Distinguished-Name* nel comando seguente con `*.mytestcluster.67281x.c4.kafka.us-east-1.amazonaws.com`. Per informazioni su come ottenere i broker bootstrap, consulta the section called “Recupero dei broker di bootstrap”.

   ```
   <path-to-your-kafka-installation>/bin/kafka-acls.sh --authorizer-properties
    --bootstrap-server BootstrapServerString --add --allow-principal
    "User:CN=Distinguished-Name" --operation Read --group=* --topic Topic-Name
   ```

2. Per concedere l'accesso in lettura a un argomento, eseguire il comando seguente sul computer client. Se utilizzi l'autenticazione TLS reciproca, utilizza lo stesso *Distinguished-Name* usato al momento della creazione della chiave privata.

   ```
   <path-to-your-kafka-installation>/bin/kafka-acls.sh --authorizer-properties
    --bootstrap-server BootstrapServerString --add --allow-principal
    "User:CN=Distinguished-Name" --operation Read --group=* --topic Topic-Name
   ```

   Per rimuovere l'accesso in lettura, è possibile eseguire lo stesso comando, sostituendo `--add` con `--remove`.

3. Per concedere l'accesso in scrittura a un argomento, eseguire il comando seguente sul computer client. Se utilizzi l'autenticazione TLS reciproca, utilizza lo stesso *Distinguished-Name* usato al momento della creazione della chiave privata.

```
<path-to-your-kafka-installation>/bin/kafka-acls.sh --authorizer-properties
 --bootstrap-server BootstrapServerString --add --allow-principal
 "User:CN=Distinguished-Name" --operation Write --topic Topic-Name
```

   Per rimuovere l'accesso in scrittura, è possibile eseguire lo stesso comando, sostituendo `--add` con `--remove`.

Modifica del gruppo di sicurezza di un cluster Amazon MSK

Questa pagina spiega come modificare il gruppo di sicurezza di un cluster MSK esistente. Potrebbe essere necessario modificare il gruppo di sicurezza di un cluster per fornire l'accesso a un determinato gruppo di utenti o per limitare l'accesso al cluster. Per ulteriori informazioni sui gruppi di sicurezza, consulta la pagina Security groups for your VPC nella Guida per l'utente di Amazon VPC.

1. Usa l'ListNodesAPI o il comando list-nodes in per AWS CLI ottenere un elenco dei broker del tuo cluster. I risultati di questa operazione includono gli ID delle interfacce di rete elastica (ENI) associate ai broker.
2. Accedi AWS Management Console e apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
3. Utilizzando l'elenco a discesa nell'angolo in alto a destra della schermata, seleziona la regione in cui è implementato il cluster.
4. Nel riquadro a sinistra, in Rete e sicurezza, scegli Interfacce di rete.
5. Seleziona il primo ENI che hai ottenuto nel primo passaggio. Scegli il menu Operazioni nella parte superiore dello schermo, quindi scegli Modifica gruppi di sicurezza. Assegna il nuovo gruppo di sicurezza a questo ENI. Ripeti questo passaggio per ciascuno degli ENI ottenuti nel primo passaggio.

> **Note**
>
> Le modifiche apportate al gruppo di sicurezza di un cluster utilizzando la console Amazon EC2 non si riflettono nelle Impostazioni di rete della console MSK.

6. Configura le regole del nuovo gruppo di sicurezza per garantire che i tuoi client abbiano accesso ai broker. Per informazioni sull'impostazione delle regole del gruppi di sicurezza, consulta la pagina Adding, Removing, and Updating Rules nella guida per l'utente di Amazon VPC.

> **Important**
>
> Se modifichi il gruppo di sicurezza associato ai broker di un cluster e poi aggiungi nuovi broker a tale cluster, Amazon MSK associa i nuovi broker al gruppo di sicurezza originale associato al cluster al momento della creazione dello stesso. Tuttavia, affinché un cluster funzioni correttamente, tutti i relativi broker devono essere associati allo stesso gruppo di sicurezza. Pertanto, se si aggiungono nuovi broker dopo aver modificato il gruppo di sicurezza, è necessario seguire nuovamente la procedura precedente e aggiornare gli ENI dei nuovi broker.

Controllo dell'accesso ad Apache ZooKeeper

Per motivi di sicurezza, puoi limitare l'accesso ai ZooKeeper nodi Apache che fanno parte del tuo cluster Amazon MSK. Per limitare l'accesso ai nodi, puoi assegnare loro un gruppo di sicurezza separato. Puoi quindi stabilire chi ottiene l'accesso a tale gruppo di sicurezza.

> **Important**
>
> Questa sezione non si applica ai cluster in esecuzione in modalità KRAFT. Per informazioni, consulta the section called “modalità KRAFT ”.

Questo argomento contiene le sezioni seguenti:

- Per collocare i ZooKeeper nodi Apache in un gruppo di sicurezza separato
- Utilizzo della sicurezza TLS con Apache ZooKeeper

Per collocare i ZooKeeper nodi Apache in un gruppo di sicurezza separato

1. Ottieni la stringa di ZooKeeper connessione Apache per il tuo cluster. Per scoprire come, consulta the section called “ZooKeeper modalità”. La stringa di connessione contiene i nomi DNS dei tuoi nodi ZooKeeper Apache.

2. Utilizzare uno strumento simile a `host` o `ping` per convertire i nomi DNS ottenuti nel passaggio precedente in indirizzi IP. Salvare questi indirizzi IP perché saranno necessari più avanti in questa procedura.
3. Accedi AWS Management Console e apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
4. Nel riquadro di navigazione, in NETWORK & SECURITY (Rete e sicurezza), scegliere Network Interfaces (Interfacce di rete).
5. Nel campo di ricerca sopra la tabella delle interfacce di rete, digitare il nome del cluster, quindi premere Invio. Questo limita il numero di interfacce di rete visualizzate nella tabella a quelle associate al cluster.
6. Selezionare la casella di controllo all'inizio della riga corrispondente alla prima interfaccia di rete nell'elenco.
7. Nel riquadro dei dettagli nella parte inferiore della pagina, cercare Primary private IPv4 IP (IP privato primario IPv4). Se questo indirizzo IP corrisponde a uno degli indirizzi IP ottenuti nel primo passaggio di questa procedura, significa che l'interfaccia di rete è assegnata a un ZooKeeper nodo Apache che fa parte del cluster. In caso contrario, deselezionare la casella di controllo accanto a questa interfaccia di rete e selezionare l'interfaccia di rete successiva nell'elenco. L'ordine di selezione delle interfacce di rete non ha importanza. Nei passaggi successivi, eseguirai le stesse operazioni su tutte le interfacce di rete assegnate ai ZooKeeper nodi Apache, una per una.
8. Quando selezioni un'interfaccia di rete che corrisponde a un ZooKeeper nodo Apache, scegli il menu Azioni nella parte superiore della pagina, quindi scegli Cambia gruppi di sicurezza. Assegnare un nuovo gruppo di sicurezza a questa interfaccia di rete. Per ulteriori informazioni sulla creazione dei gruppi di sicurezza, consulta la pagina Creating a Security Group nella documentazione di Amazon VPC.
9. Ripeti il passaggio precedente per assegnare lo stesso nuovo gruppo di sicurezza a tutte le interfacce di rete associate ai ZooKeeper nodi Apache del cluster.
10. È ora possibile scegliere chi dispone dell'accesso a questo nuovo gruppo di sicurezza. Per informazioni sull'impostazione delle regole dei gruppi di sicurezza, consulta la pagina Adding, Removing, and Updating Rules nella documentazione di Amazon VPC.

Utilizzo della sicurezza TLS con Apache ZooKeeper

Puoi utilizzare la sicurezza TLS per la crittografia in transito tra i tuoi client e i tuoi nodi Apache. ZooKeeper Per implementare la sicurezza TLS con i tuoi ZooKeeper nodi Apache, procedi come segue:

- I cluster devono utilizzare Apache Kafka versione 2.5.1 o successiva per utilizzare la sicurezza TLS con Apache. ZooKeeper
- Abilita la sicurezza TLS quando crei o configuri il cluster. I cluster creati con Apache Kafka versione 2.5.1 o successiva con TLS abilitato utilizzano automaticamente la sicurezza TLS con gli endpoint Apache. ZooKeeper Per informazioni sulla configurazione della sicurezza TLS, consulta la pagina Quali sono i primi passi per iniziare a utilizzare la crittografia?.
- Recupera gli endpoint TLS Apache utilizzando l'operazione. ZooKeeper DescribeCluster
- Crea un file di ZooKeeper configurazione Apache da utilizzare con `kafka-acls.sh`gli strumenti `kafka-configs.sh` and o con la shell. ZooKeeper Con ogni strumento, si utilizza il `--zk-tls-config-file` parametro per specificare la configurazione di Apache ZooKeeper .

  L'esempio seguente mostra un tipico file di configurazione di Apache ZooKeeper :

  ```
  zookeeper.ssl.client.enable=true
  zookeeper.clientCnxnSocket=org.apache.zookeeper.ClientCnxnSocketNetty
  zookeeper.ssl.keystore.location=kafka.jks
  zookeeper.ssl.keystore.password=test1234
  zookeeper.ssl.truststore.location=truststore.jks
  zookeeper.ssl.truststore.password=test1234
  ```

- Per altri comandi (come`kafka-topics`), è necessario utilizzare la variabile di `KAFKA_OPTS` ambiente per configurare i parametri di Apache ZooKeeper. L'esempio seguente mostra come configurare la variabile di `KAFKA_OPTS` ambiente per passare i ZooKeeper parametri Apache ad altri comandi:

  ```
  export KAFKA_OPTS="
  -Dzookeeper.clientCnxnSocket=org.apache.zookeeper.ClientCnxnSocketNetty
  -Dzookeeper.client.secure=true
  -Dzookeeper.ssl.trustStore.location=/home/ec2-user/kafka.client.truststore.jks
  -Dzookeeper.ssl.trustStore.password=changeit"
  ```

Dopo aver configurato la variabile di ambiente KAFKA_0PTS, è possibile utilizzare normalmente i comandi della CLI. L'esempio seguente crea un argomento di Apache Kafka utilizzando la ZooKeeper configurazione di Apache dalla variabile di ambiente: KAFKA_0PTS

```
<path-to-your-kafka-installation>/bin/kafka-topics.sh --create --
zookeeper ZooKeeperTLSConnectString --replication-factor 3 --partitions 1 --topic
 AWSKafkaTutorialTopic
```

> **Note**
>
> I nomi dei parametri utilizzati nel file di ZooKeeper configurazione di Apache e quelli utilizzati nella variabile di KAFKA_0PTS ambiente non sono coerenti. Presta attenzione ai nomi che usi con ciascun parametri nel file di configurazione e nella variabile di ambiente KAFKA_0PTS.

Per ulteriori informazioni sull'accesso ai ZooKeeper nodi Apache con TLS, vedi KIP-515: Abilita il client ZK a usare la nuova autenticazione supportata da TLS.

Registrazione

Puoi inviare i log del broker Apache Kafka a uno o più dei seguenti tipi di destinazione: Amazon Logs, Amazon S3 CloudWatch , Amazon Data Firehose. Puoi anche registrare le chiamate API Amazon MSK con AWS CloudTrail.

Log di broker

I log di broker consentono di risolvere i problemi delle applicazioni Apache Kafka e di analizzare le comunicazioni con il cluster MSK. È possibile configurare il cluster MSK nuovo o esistente per fornire i log dei broker a livello Info a uno o più dei seguenti tipi di risorse di destinazione: un gruppo di CloudWatch log, un bucket S3, un flusso di distribuzione Firehose. Tramite Firehose è quindi possibile inviare i dati di registro dal flusso di distribuzione a OpenSearch Service. È necessario creare una risorsa di destinazione prima di configurare il cluster per consegnargli i log del broker. Amazon MSK non crea queste risorse di destinazione se non esistono già. Per informazioni su questi tre tipi di risorse di destinazione e su come crearle, consultare la documentazione seguente:

- CloudWatch Registri Amazon

- Amazon S3
- Amazon Data Firehose

Autorizzazioni richieste

Per configurare una destinazione per i log del broker Amazon MSK, l'identità IAM che utilizzi per le operazioni Amazon MSK deve disporre delle autorizzazioni descritte nella policy AWS politica gestita: AmazonMSK FullAccess.

Per eseguire lo streaming dei log di broker a un bucket S3, è richiesta anche l'autorizzazione `s3:PutBucketPolicy`. Per informazioni sulle policy dei bucket S3, consulta la pagina How Do I Add an S3 Bucket Policy? nella Guida per l'utente di Amazon S3. Per informazioni sulle policy IAM in generale, consulta la pagina Access Management nella Guida per l'utente di IAM.

Policy della chiave KMS necessaria per l'utilizzo con i bucket SSE-KMS

Se hai abilitato la crittografia lato server per il tuo bucket S3 utilizzando chiavi AWS KMS gestite (SSE-KMS) con una chiave gestita dal cliente, aggiungi quanto segue alla policy chiave per la tua chiave KMS in modo che Amazon MSK possa scrivere i file del broker nel bucket.

```
{
  "Sid": "Allow Amazon MSK to use the key.",
  "Effect": "Allow",
  "Principal": {
    "Service": [
      "delivery.logs.amazonaws.com"
    ]
  },
  "Action": [
    "kms:Encrypt",
    "kms:Decrypt",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*",
    "kms:DescribeKey"
  ],
  "Resource": "*"
}
```

Configurazione dei log del broker utilizzando AWS Management Console

Se stai creando un nuovo cluster, cerca l'intestazione Broker log delivery (Recapito del log del broker) nella sezione Monitoring (Monitoraggio) . Puoi specificare le destinazioni a cui Amazon MSK deve consegnare i log del broker.

Per un cluster esistente, scegli il cluster dall'elenco di cluster, quindi seleziona la scheda Proprietà. Scorri verso il basso fino alla sezione Consegna dei log e scegli il relativo pulsante Modifica. Puoi specificare le destinazioni a cui Amazon MSK deve consegnare i log del broker.

Configurazione dei log del broker utilizzando il AWS CLI

Quando utilizzi i comandi `create-cluster` o `update-monitoring`, puoi specificare facoltativamente il parametro `logging-info` e passarlo a una struttura JSON come nell'esempio seguente. In questo JSON, tutti e tre i tipi di destinazione sono facoltativi.

```
{
  "BrokerLogs": {
    "S3": {
      "Bucket": "ExampleBucketName",
      "Prefix": "ExamplePrefix",
      "Enabled": true
    },
    "Firehose": {
      "DeliveryStream": "ExampleDeliveryStreamName",
      "Enabled": true
    },
    "CloudWatchLogs": {
      "Enabled": true,
      "LogGroup": "ExampleLogGroupName"
    }
  }
}
```

Configurazione dei log del broker mediante l'API

È possibile specificare la `loggingInfo` struttura opzionale nel file JSON che si passa alle operazioni or. CreateClusterUpdateMonitoring

> **Note**
>
> Per impostazione predefinita, quando la registrazione del broker è abilitata, Amazon MSK registra i log di livello `INFO` nelle destinazioni specificate. Tuttavia, gli utenti di Apache Kafka 2.4.X e versioni successive possono impostare dinamicamente il livello di log del broker su uno qualsiasi dei livelli di log log4j. Per informazioni sull'impostazione dinamica del livello di log del broker, consulta la pagina KIP-412: Extend Admin API to support dynamic application log levels. Se imposti dinamicamente il livello di registro su `DEBUG` o`TRACE`, ti consigliamo di utilizzare Amazon S3 o Firehose come destinazione del log. Se utilizzi CloudWatch Logs come destinazione di log e abiliti `DEBUG` o `TRACE` livelli dinamicamente la registrazione, Amazon MSK può fornire continuamente un campione di log. Ciò può influire in modo significativo sulle prestazioni del broker e deve essere utilizzato solo quando il livello di log `INFO` non è sufficientemente dettagliato da consentire di determinare la causa principale di un problema.

Registrazione delle chiamate API di AWS CloudTrail con

> **Note**
>
> AWS CloudTrail i log sono disponibili per Amazon MSK solo quando li usi. Controllo degli accessi IAM

Amazon MSK è integrato con AWS CloudTrail, un servizio che fornisce un registro delle azioni intraprese da un utente, un ruolo o un AWS servizio in Amazon MSK. CloudTrail acquisisce le chiamate API come eventi. Le chiamate acquisite includono le chiamate dalla console Amazon MSK e le chiamate di codice alle operazioni API di Amazon MSK. Registra anche le operazioni di Apache Kafka come la creazione e la modifica di argomenti e gruppi.

Se crei un trail, puoi abilitare la distribuzione continua di CloudTrail eventi a un bucket Amazon S3, inclusi gli eventi per Amazon MSK. Se non configuri un percorso, puoi comunque visualizzare gli eventi più recenti nella CloudTrail console nella cronologia degli eventi. Utilizzando le informazioni raccolte da CloudTrail, puoi determinare la richiesta effettuata ad Amazon MSK o l'azione Apache Kafka, l'indirizzo IP da cui è stata effettuata la richiesta, chi ha effettuato la richiesta, quando è stata effettuata e dettagli aggiuntivi.

Per ulteriori informazioni CloudTrail, incluso come configurarlo e abilitarlo, consulta la Guida per l'utente.AWS CloudTrail

Informazioni su Amazon MSK in CloudTrail

CloudTrail è abilitato sul tuo account Amazon Web Services al momento della creazione dell'account. Quando si verifica un'attività di evento supportata in un cluster MSK, tale attività viene registrata in un CloudTrail evento insieme ad altri eventi di AWS servizio nella cronologia degli eventi. Puoi visualizzare, cercare e scaricare gli eventi recenti nell'account Amazon Web Services. Per ulteriori informazioni, vedere Visualizzazione degli eventi con la cronologia degli CloudTrail eventi.

Per una registrazione continua degli eventi nell'account Amazon Web Services che includa gli eventi per Amazon MSK, crea un percorso. Un trail consente di CloudTrail inviare file di log a un bucket Amazon S3. Per impostazione predefinita, quando si crea un trail nella console, il trail sarà valido in tutte le Regioni . Il trail registra gli eventi di tutte le Regioni nella partizione AWS e distribuisce i file di log nel bucket Amazon S3 specificato. Inoltre, puoi configurare altri servizi Amazon per analizzare ulteriormente e agire in base ai dati sugli eventi raccolti nei CloudTrail log. Per ulteriori informazioni, consulta gli argomenti seguenti:

- Panoramica della creazione di un trail
- CloudTrail Servizi e integrazioni supportati
- Configurazione delle notifiche Amazon SNS per CloudTrail
- Ricezione di file di CloudTrail registro da più regioni e ricezione di file di CloudTrail registro da più account

Amazon MSK registra tutte le operazioni di Amazon MSK come eventi nei CloudTrail file di registro. Inoltre, registra le seguenti operazioni di Apache Kafka.

- cluster kafka: DescribeClusterDynamicConfiguration
- ammasso kafka: AlterClusterDynamicConfiguration
- ammasso kafka: CreateTopic
- ammasso kafka: DescribeTopicDynamicConfiguration
- ammasso kafka: AlterTopic
- ammasso kafka: AlterTopicDynamicConfiguration
- ammasso kafka: DeleteTopic

Ogni evento o voce di log contiene informazioni sull'utente che ha generato la richiesta. Le informazioni di identità consentono di determinare quanto segue:

- Se la richiesta è stata effettuata con credenziali utente root o utente AWS Identity and Access Management (IAM).
- Se la richiesta è stata effettuata con le credenziali di sicurezza temporanee per un ruolo o un utente federato.
- Se la richiesta è stata effettuata da un altro AWS servizio.

Per ulteriori informazioni, vedete l'elemento CloudTrail userIdentity.

Esempio: voci del file di log di Amazon MSK

Un trail è una configurazione che consente la distribuzione di eventi come file di log in un bucket Amazon S3 specificato dall'utente. CloudTrail i file di registro contengono una o più voci di registro. Un evento rappresenta una singola richiesta da un'fonte e include informazioni sull'azione richiesta, data e ora dell'azione, parametri richiesti e così via. CloudTrail i file di registro non sono una traccia ordinata delle chiamate API pubbliche e delle azioni di Apache Kafka, quindi non vengono visualizzati in un ordine specifico.

L'esempio seguente mostra le voci di CloudTrail registro che illustrano le azioni `DescribeCluster` e `DeleteCluster` Amazon MSK.


```
{
  "Records": [
    {
      "eventVersion": "1.05",
      "userIdentity": {
        "type": "IAMUser",
        "principalId": "ABCDEF0123456789ABCDE",
        "arn": "arn:aws:iam::012345678901:user/Joe",
        "accountId": "012345678901",
        "accessKeyId": "AIDACKCEVSQ6C2EXAMPLE",
        "userName": "Joe"
      },
      "eventTime": "2018-12-12T02:29:24Z",
      "eventSource": "kafka.amazonaws.com",
      "eventName": "DescribeCluster",
      "awsRegion": "us-east-1",
      "sourceIPAddress": "192.0.2.0",
```

```
      "userAgent": "aws-cli/1.14.67 Python/3.6.0 Windows/10 botocore/1.9.20",
      "requestParameters": {
        "clusterArn": "arn%3Aaws%3Akafka%3Aus-east-1%3A012345678901%3Acluster
%2Fexamplecluster%2F01234567-abcd-0123-abcd-abcd0123efa-2"
      },
      "responseElements": null,
      "requestID": "bd83f636-fdb5-abcd-0123-157e2fbf2bde",
      "eventID": "60052aba-0123-4511-bcde-3e18dbd42aa4",
      "readOnly": true,
      "eventType": "AwsApiCall",
      "recipientAccountId": "012345678901"
    },
    {
      "eventVersion": "1.05",
      "userIdentity": {
        "type": "IAMUser",
        "principalId": "ABCDEF0123456789ABCDE",
        "arn": "arn:aws:iam::012345678901:user/Joe",
        "accountId": "012345678901",
        "accessKeyId": "AIDACKCEVSQ6C2EXAMPLE",
        "userName": "Joe"
      },
      "eventTime": "2018-12-12T02:29:40Z",
      "eventSource": "kafka.amazonaws.com",
      "eventName": "DeleteCluster",
      "awsRegion": "us-east-1",
      "sourceIPAddress": "192.0.2.0",
      "userAgent": "aws-cli/1.14.67 Python/3.6.0 Windows/10 botocore/1.9.20",
      "requestParameters": {
        "clusterArn": "arn%3Aaws%3Akafka%3Aus-east-1%3A012345678901%3Acluster
%2Fexamplecluster%2F01234567-abcd-0123-abcd-abcd0123efa-2"
      },
      "responseElements": {
        "clusterArn": "arn:aws:kafka:us-east-1:012345678901:cluster/
examplecluster/01234567-abcd-0123-abcd-abcd0123efa-2",
        "state": "DELETING"
      },
      "requestID": "c6bfb3f7-abcd-0123-afa5-293519897703",
      "eventID": "8a7f1fcf-0123-abcd-9bdb-1ebf0663a75c",
      "readOnly": false,
      "eventType": "AwsApiCall",
      "recipientAccountId": "012345678901"
    }
  ]
```

```
}
```

L'esempio seguente mostra una voce di CloudTrail registro che illustra l'`kafka-cluster:CreateTopic`azione.

```
{
  "eventVersion": "1.08",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "ABCDEFGH1IJKLMN2P34Q5",
    "arn": "arn:aws:iam::111122223333:user/Admin",
    "accountId": "111122223333",
    "accessKeyId": "CDEFAB1C2UUUUU3AB4TT",
    "userName": "Admin"
  },
  "eventTime": "2021-03-01T12:51:19Z",
  "eventSource": "kafka-cluster.amazonaws.com",
  "eventName": "CreateTopic",
  "awsRegion": "us-east-1",
  "sourceIPAddress": "198.51.100.0/24",
  "userAgent": "aws-msk-iam-auth/unknown-version/aws-internal/3 aws-sdk-java/1.11.970
 Linux/4.14.214-160.339.amzn2.x86_64 OpenJDK_64-Bit_Server_VM/25.272-b10 java/1.8.0_272
 scala/2.12.8 vendor/Red_Hat,_Inc.",
  "requestParameters": {
    "kafkaAPI": "CreateTopics",
    "resourceARN": "arn:aws:kafka:us-east-1:111122223333:topic/IamAuthCluster/3ebafd8e-
dae9-440d-85db-4ef52679674d-1/Topic9"
  },
  "responseElements": null,
  "requestID": "e7c5e49f-6aac-4c9a-a1d1-c2c46599f5e4",
  "eventID": "be1f93fd-4f14-4634-ab02-b5a79cb833d2",
  "readOnly": false,
  "eventType": "AwsApiCall",
  "managementEvent": true,
  "eventCategory": "Management",
  "recipientAccountId": "111122223333"
}
```

Convalida della conformità per Streaming gestito da Amazon per Apache Kafka

Revisori di terze parti valutano la sicurezza e la conformità di Streaming gestito da Amazon per Apache Kafka nell'ambito dei programmi di conformità di AWS . Questi includono PCI e HIPAA BAA.

Per un elenco di AWS servizi nell'ambito di programmi di conformità specifici, consulta Amazon Services in Scope by Compliance Program . Per informazioni generali, consulta Programmi di AWS conformità Programmi di .

È possibile scaricare report di audit di terze parti utilizzando AWS Artifact. Per ulteriori informazioni, consulta Scaricamento dei report in AWS Artifact .

La tua responsabilità di conformità quando usi Amazon MSK è determinata dalla sensibilità dei tuoi dati, dagli obiettivi di conformità della tua azienda e dalle leggi e dai regolamenti applicabili. AWS fornisce le seguenti risorse per contribuire alla conformità:

- Security and Compliance Quick Start Guides (Guide Quick Start Sicurezza e compliance): queste guide alla distribuzione illustrano considerazioni relative all'architettura e forniscono procedure per la distribuzione di ambienti di base incentrati sulla sicurezza e sulla conformità su AWS.
- Whitepaper sull'architettura per la sicurezza e la conformità HIPAA: questo white paper descrive come le aziende possono utilizzare per creare applicazioni conformi allo standard HIPAA. AWS
- AWS Risorse per la conformità Risorse per la conformità: questa raccolta di potrebbe riguardare il settore e la località in cui operate.
- Valutazione delle risorse con le regole nella Guida per gli AWS Config sviluppatori: il AWS Config servizio valuta la conformità delle configurazioni delle risorse alle pratiche interne, alle linee guida del settore e alle normative.
- AWS Security Hub— Questo AWS servizio offre una visione completa dello stato di sicurezza dell'utente, AWS che consente di verificare la conformità agli standard e alle best practice del settore della sicurezza.

Resilienza in Streaming gestito da Amazon per Apache Kafka

L'infrastruttura AWS globale è costruita attorno a regioni e zone di disponibilità. AWS AWS Le regioni forniscono più zone di disponibilità fisicamente separate e isolate, collegate con reti a bassa latenza, ad alto throughput e altamente ridondanti. Con le zone di disponibilità, puoi progettare e gestire

applicazioni e database che eseguono automaticamente il failover tra zone di disponibilità senza interruzioni. Le zone di disponibilità sono più disponibili, tolleranti ai guasti e scalabili rispetto alle infrastrutture a data center singolo o multiplo tradizionali.

Per ulteriori informazioni su AWS regioni e zone di disponibilità, consulta Global Infrastructure.AWS

Sicurezza dell'infrastruttura in Streaming gestito da Amazon per Apache Kafka

In quanto servizio gestito, Amazon Managed Streaming for Apache Kafka è protetto AWS dalle procedure di sicurezza di rete globali descritte nel white paper di Amazon Web Services: Overview of Security Processes.

Utilizzi chiamate API AWS pubblicate per accedere ad Amazon MSK attraverso la rete. I client devono supportare Transport Layer Security (TLS) 1.0 o versioni successive. È consigliabile TLS 1.2 o versioni successive. I client devono, inoltre, supportare le suite di cifratura con PFS (Perfect Forward Secrecy), ad esempio Ephemeral Diffie-Hellman (DHE) o Elliptic Curve Ephemeral Diffie-Hellman (ECDHE). La maggior parte dei sistemi moderni, come Java 7 e versioni successive, supporta tali modalità.

Inoltre, le richieste devono essere firmate utilizzando un ID chiave di accesso e una chiave di accesso segreta associata a un principale IAM. O puoi utilizzare AWS Security Token Service (AWS STS) per generare credenziali di sicurezza temporanee per sottoscrivere le richieste.

Connessione a un cluster Amazon MSK

Per impostazione predefinita, i client possono accedere a un cluster MSK solo se si trovano nello stesso VPC del cluster. Per impostazione predefinita, tutte le comunicazioni tra i client Kafka e il cluster MSK sono private e i dati di streaming non attraversano mai Internet. Per connetterti al cluster MSK da un client che si trova nello stesso VPC del cluster, assicurati che il gruppo di sicurezza del cluster disponga di una regola in entrata che accetti il traffico dal gruppo di sicurezza del client. Per informazioni sull'impostazione di queste regole, consulta Regole del gruppo di sicurezza. Per un esempio di come accedere a un cluster da un'istanza Amazon EC2 che si trova nello stesso VPC del cluster, consulta la pagina Nozioni di base.

Per connetterti al tuo cluster MSK da un client esterno al VPC del cluster, vedi Accesso dall' AWS interno ma dall'esterno del VPC del cluster.

Argomenti

- Accesso pubblico
- Accesso dall'interno AWS ma dall'esterno del VPC del cluster

Accesso pubblico

Amazon MSK ti offre la possibilità di attivare l'accesso pubblico ai broker dei cluster MSK che eseguono Apache Kafka 2.6.0 o versioni successive. Per motivi di sicurezza, non è possibile attivare l'accesso pubblico durante la creazione di un cluster MSK. Tuttavia, è possibile aggiornare un cluster esistente per renderlo accessibile al pubblico. È inoltre possibile creare un nuovo cluster e aggiornarlo in modo da renderlo accessibile al pubblico.

È possibile attivare l'accesso pubblico a un cluster MSK senza costi aggiuntivi, ma per il trasferimento AWS dei dati in entrata e in uscita dal cluster si applicano i costi standard di trasferimento dei dati. Per ulteriori informazioni, consulta la pagina Prezzi di Amazon EC2 on demand.

Per attivare l'accesso pubblico a un cluster, assicurati innanzitutto che il cluster soddisfi tutte le seguenti condizioni:

- Le sottoreti associate al cluster devono essere pubbliche. Ciò significa che le sottoreti devono avere una tabella di routing associata a un gateway Internet. Per ulteriori informazioni sulla creazione e il collegamento di un gateway Internet, consulta la pagina Internet gateways nella Guida per l'utente di Amazon VPC.

- Il controllo degli accessi non autenticati deve essere disattivato e almeno uno dei seguenti metodi di controllo degli accessi deve essere attivo: SASL/IAM, SASL/SCRAM, mTLS. Per informazioni su come aggiornare il metodo di controllo degli accessi di un cluster, consulta la pagina the section called “Aggiornamento della sicurezza”.
- La crittografia all'interno del cluster deve essere attiva. Per impostazione predefinita durante la creazione di un cluster, la crittografia è attiva. Non è possibile attivare la crittografia all'interno del cluster se esso è stato creato con questa opzione disattivata. Pertanto, non è possibile attivare l'accesso pubblico per il cluster se esso è stato creato con la crittografia all'interno del cluster disattivata.
- Il traffico non crittografato tra broker e client deve essere disattivato. Per informazioni su come disattivarlo se è attivato, consulta la pagina the section called “Aggiornamento della sicurezza”.
- Se si utilizzano i metodi di controllo degli accessi SASL/SCRAM o mTLS, è necessario impostare le ACL di Apache Kafka per il cluster. Dopo avere impostato le ACL di Apache Kafka per il cluster, aggiorna la configurazione del cluster in modo che la proprietà `allow.everyone.if.no.acl.found` del cluster sia impostata su false. Per informazioni su come aggiornare la configurazione di un cluster, consulta la pagina the section called “Operazioni di configurazione”. Se utilizzi il Controllo degli accessi IAM e desideri applicare policy di autorizzazione o aggiornare le tue policy esistenti, consulta la sezione the section called “Controllo degli accessi IAM”. Per ulteriori informazioni sulle ACL di Apache Kafka, consulta la pagina the section called “ACL Apache Kafka”.

Dopo esserti assicurato che un cluster MSK soddisfi le condizioni sopra elencate AWS Management Console, puoi utilizzare l' AWS CLI API Amazon MSK per attivare l'accesso pubblico. Dopo aver attivato l'accesso pubblico a un cluster, puoi recuperare una stringa bootstrap-brokers pubblica relativa al cluster. Per informazioni su come recuperare i broker di bootstrap per un cluster, consulta la pagina the section called “Recupero dei broker di bootstrap”.

> **Important**
>
> Oltre ad attivare l'accesso pubblico, assicurati che i gruppi di sicurezza del cluster dispongano di regole TCP in entrata che consentano l'accesso pubblico dal tuo indirizzo IP. Ti consigliamo di impostare tali regole di modo che siano il più restrittive possibile. Per ulteriori informazioni sui gruppi di sicurezza e le regole in entrata, consulta la pagina Security groups for your VPC nella Guida per l'utente di Amazon VPC. Per i numeri di porta, consulta la pagina the section called “Informazioni sulle porte”. Per istruzioni su come modificare il

gruppo di sicurezza di un cluster, consulta la pagina the section called “Modifica dei gruppi di sicurezza”.

> **Note**
>
> Se dopo avere seguito le istruzioni seguenti per attivare l'accesso pubblico non riesci comunque ad accedere al cluster, consulta la pagina the section called “Impossibile accedere al cluster con accesso pubblico attivato”.

Attivazione dell'accesso pubblico tramite la console

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Nell'elenco dei cluster, scegli quello per il quale attivare l'accesso pubblico.
3. Scegli la scheda Proprietà, quindi trova la sezione Impostazioni di rete.
4. Scegli Modifica accesso pubblico.

Attivazione dell'accesso pubblico tramite AWS CLI

1. Esegui il AWS CLI comando seguente, sostituendo *Current-Cluster-Version con l'ARN ClusterArne la versione* corrente del cluster. Per trovare la versione corrente del cluster, utilizzare l'operazione o il comando describe-cluster. DescribeCluster AWS CLI Una versione di esempio è `KTVPDKIKX0DER`.

   ```
   aws kafka update-connectivity --cluster-arn ClusterArn --current-
   version Current-Cluster-Version --connectivity-info '{"PublicAccess": {"Type":
    "SERVICE_PROVIDED_EIPS"}}'
   ```

   L'output di questo comando `update-connectivity` è simile all'esempio JSON seguente.

   ```
   {
       "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/exampleClusterName/
   abcdefab-1234-abcd-5678-cdef0123ab01-2",
       "ClusterOperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
   operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
   abcd-4f7f-1234-9876543210ef"
   ```

```
}
```

> **Note**
>
> Per disattivare l'accesso pubblico, usa un AWS CLI comando simile, ma con le seguenti informazioni di connettività:
>
> ```
> '{"PublicAccess": {"Type": "DISABLED"}}'
> ```

2. Per ottenere il risultato dell'`update-connectivity`operazione, esegui il comando seguente, sostituendo *`ClusterOperationArn`* con l'ARN ottenuto nell'output del `update-connectivity` comando.

```
aws kafka describe-cluster-operation --cluster-operation-arn ClusterOperationArn
```

L'output di questo comando `describe-cluster-operation` è simile all'esempio JSON seguente.

```
{
    "ClusterOperationInfo": {
        "ClientRequestId": "982168a3-939f-11e9-8a62-538df00285db",
        "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/
exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2",
        "CreationTime": "2019-06-20T21:08:57.735Z",
        "OperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef",
        "OperationState": "UPDATE_COMPLETE",
        "OperationType": "UPDATE_CONNECTIVITY",
        "SourceClusterInfo": {
            "ConnectivityInfo": {
                "PublicAccess": {
                    "Type": "DISABLED"
                }
            }
        },
        "TargetClusterInfo": {
            "ConnectivityInfo": {
                "PublicAccess": {
                    "Type": "SERVICE_PROVIDED_EIPS"
```

```
                }
            }
        }
    }
}
```

Se il valore di `OperationState` è `UPDATE_IN_PROGRESS`, attendi qualche minuto, quindi esegui nuovamente il comando `describe-cluster-operation`.

Attivazione dell'accesso pubblico tramite l'API di Amazon MSK

- Per utilizzare l'API per attivare o disattivare l'accesso pubblico a un cluster, consulta. UpdateConnectivity

> **Note**
>
> Per motivi di sicurezza, Amazon MSK non consente l'accesso pubblico ai nodi di controllo Apache ZooKeeper o KRAFT.

Accesso dall'interno AWS ma dall'esterno del VPC del cluster

Per connettersi a un cluster MSK dall'interno AWS ma dall'esterno dell'Amazon VPC del cluster, esistono le seguenti opzioni.

Peering Amazon VPC

Per connettersi al cluster MSK da un VPC diverso dal VPC del cluster, è possibile creare una connessione peering tra i due VPC. Per informazioni sul peering VPC, consulta la Guida al peering di VPC di Amazon.

AWS Direct Connect

AWS Direct Connect collega la rete locale a un cavo in fibra ottica AWS Ethernet standard da 1 o 10 gigabit. Un'estremità del cavo è collegata al router, l'altra a un router. AWS Direct Connect Con questa connessione, puoi creare interfacce virtuali direttamente sul AWS cloud e Amazon VPC, aggirando i provider di servizi Internet nel tuo percorso di rete. Per ulteriori informazioni, consulta AWS Direct Connect.

AWS Transit Gateway

AWS Transit Gateway è un servizio che ti consente di connettere i tuoi VPC e le tue reti locali a un unico gateway. Per informazioni su come utilizzare AWS Transit Gateway, consulta AWS Transit Gateway.

Connessioni VPN

Puoi connettere il VPC del cluster MSK a reti e utenti remoti utilizzando le opzioni di connettività VPN descritte nel seguente argomento: VPN Connections.

Proxy REST

Puoi installare un proxy REST in un'istanza in esecuzione all'interno dell'Amazon VPC del cluster. I proxy REST consentono ai produttori e ai consumatori di comunicare con il cluster attraverso richieste API HTTP.

Connettività multi-VPC per regioni multiple

Nel documento seguente vengono descritte le opzioni di connettività per più VPC che risiedono in diverse regioni: Multiple Region Multi-VPC Connectivity.

Connettività privata multi-VPC a regione singola

La connettività privata multi-VPC (con tecnologia AWS PrivateLink) per i cluster Amazon Managed Streaming for Apache Kafka (Amazon MSK) è una funzionalità che consente di connettere più rapidamente i client Kafka ospitati in diversi Virtual Private Cloud (VPC) e account a un cluster Amazon MSK. AWS

Consulta la sezione Single Region multi-VPC connectivity for cross-account clients.

La rete EC2-Classic è stata ritirata

Amazon MSK non supporta più le istanze Amazon EC2 in esecuzione con reti Amazon EC2-Classic.

Vedi EC2-Classic Networking is Retiring: ecco come prepararsi.

Connettività privata multi-VPC di Amazon MSK in un'unica regione

La connettività privata multi-VPC (con tecnologia AWS PrivateLink) per i cluster Amazon Managed Streaming for Apache Kafka (Amazon MSK) è una funzionalità che consente di connettere più

rapidamente i client Kafka ospitati in diversi Virtual Private Cloud (VPC) e account a un cluster Amazon MSK. AWS

La connettività privata multi-VPC è una soluzione gestita che semplifica l'infrastruttura di rete per la connettività multi-VPC e multi-account. I client possono connettersi al cluster Amazon MSK PrivateLink mantenendo tutto il traffico all'interno della AWS rete. La connettività privata multi-VPC per i cluster Amazon MSK è disponibile in tutte le regioni in AWS cui è disponibile Amazon MSK.

Argomenti

Cos'è la connettività privata multi-VPC?

La connettività privata multi-VPC per Amazon MSK è un'opzione di connettività che consente di connettere client Apache Kafka ospitati in diversi account e cloud privati virtuali (VPC) a un cluster MSK. AWS

Amazon MSK semplifica l'accesso multi-account con le policy del cluster. Queste politiche consentono al proprietario del cluster di concedere autorizzazioni ad altri AWS account per stabilire una connettività privata al cluster MSK.

Vantaggi della connettività privata multi-VPC

La connettività privata multi-VPC presenta diversi vantaggi rispetto ad altre soluzioni di connettività:

- Automatizza la gestione operativa della soluzione di connettività. AWS PrivateLink
- Consente la sovrapposizione degli IP tra i VPC connessi, eliminando la necessità di mantenere IP non sovrapposti, peering e tabelle di routing complesse associate ad altre soluzioni di connettività VPC.

Si utilizza una politica di cluster per il cluster MSK per definire quali AWS account dispongono delle autorizzazioni per configurare la connettività privata tra account al cluster MSK. L'amministratore multi-account può delegare le autorizzazioni ai ruoli o agli utenti appropriati. In combinazione con l'autenticazione del client IAM, puoi utilizzare la policy del cluster anche per definire in modo granulare le autorizzazioni del piano dati Kafka per i client che si connettono.

Requisiti e limitazioni per la connettività privata multi-VPC

Tieni conto di questi requisiti del cluster MSK per l'esecuzione della connettività privata multi-VPC:

- La connettività privata multi-VPC è supportata solo su Apache Kafka 2.7.1 o versioni successive. Assicurati che tutti i client utilizzati con il cluster MSK eseguano versioni di Apache Kafka compatibili con il cluster.
- La connettività privata multi-VPC supporta i tipi di autenticazione IAM, TLS e SASL/SCRAM. I cluster non autenticati non possono utilizzare la connettività privata multi-VPC.
- Se si utilizzano i metodi di controllo degli accessi SASL/SCRAM o mTLS, è necessario impostare le ACL di Apache Kafka per il cluster. Innanzitutto, imposta le ACL di Apache Kafka per il cluster. Quindi, aggiorna la configurazione del cluster in modo che la proprietà `allow.everyone.if.no.acl.found` sia impostata su false per il cluster. Per informazioni su come aggiornare la configurazione di un cluster, consulta la pagina the section called “Operazioni di configurazione”. Se utilizzi il Controllo degli accessi IAM e desideri applicare policy di autorizzazione o aggiornare le tue policy esistenti, consulta la sezione the section called “Controllo degli accessi IAM”. Per ulteriori informazioni sulle ACL di Apache Kafka, consulta la pagina the section called “ACL Apache Kafka”.
- La connettività privata multi-VPC non supporta il tipo di istanza t3.small.
- La connettività privata multi-VPC non è supportata in tutte AWS le regioni, ma solo negli AWS account all'interno della stessa regione.
- Amazon MSK non supporta la connettività privata multi-VPC ai nodi ZooKeeper.

Guida introduttiva all'utilizzo della connettività privata multi-VPC

Argomenti

- Passaggio 1: sul cluster MSK nell'account A, attiva la connettività multi-VPC per lo schema di autenticazione IAM sul cluster
- Passaggio 2: collegamento di una policy del cluster al cluster MSK

- Passaggio 3: operazioni dell'utente multi-account per configurare connessioni VPC gestite dal client

Questo tutorial utilizza un caso d'uso comune come esempio di come utilizzare la connettività multi-VPC per connettere privatamente un client Apache Kafka a un cluster MSK dall'interno AWS ma dall'esterno del VPC del cluster. Questo processo richiede che l'utente multi-account crei una connessione e una configurazione VPC gestite da MSK per ogni client, comprese le autorizzazioni client richieste. Il processo richiede inoltre che il proprietario del cluster MSK abiliti la PrivateLink connettività sul cluster MSK e selezioni gli schemi di autenticazione per controllare l'accesso al cluster.

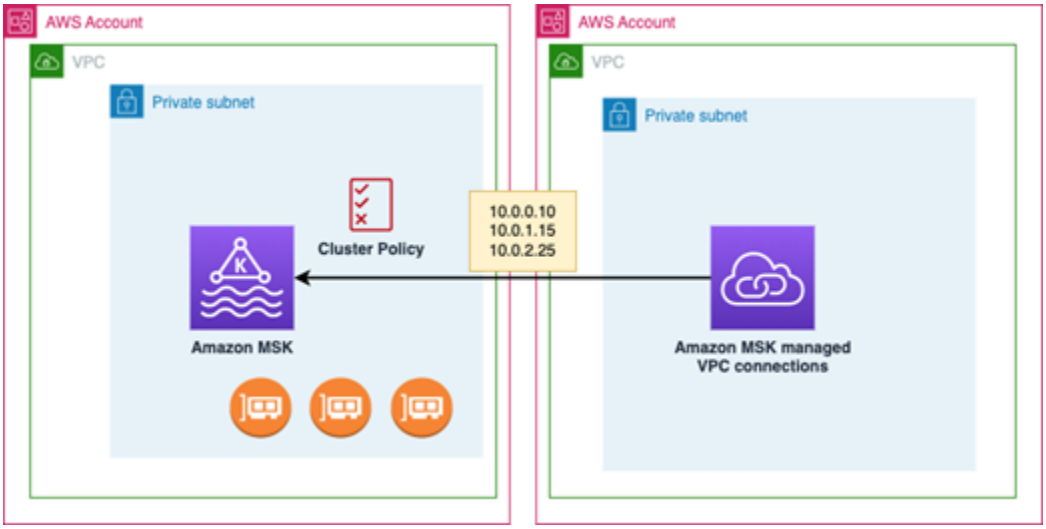
In diverse parti di questo tutorial, scegliamo le opzioni che si applicano a questo esempio. Ciò non significa che siano le uniche opzioni che funzionano per la configurazione di un cluster MSK o delle istanze client.

La configurazione di rete per questo caso d'uso è la seguente:

- Un utente multi-account (client Kafka) e un cluster MSK si trovano nella stessa rete/regione AWS , ma in account diversi:
  - Cluster MSK nell'account A
  - Cliente Kafka nell'account B
- L'utente multi-account si connetterà privatamente al cluster MSK utilizzando lo schema di autenticazione IAM.

Questo tutorial presuppone che esista un cluster MSK assegnato creato con Apache Kafka versione 2.7.1 o successiva. Il cluster MSK deve essere in uno stato ACTIVE prima di iniziare il processo di configurazione. Per evitare potenziali perdite di dati o tempi di inattività, i client che utilizzeranno una connessione privata multi-VPC per connettersi al cluster devono utilizzare versioni di Apache Kafka compatibili con il cluster.

Il diagramma seguente illustra l'architettura della connettività multi-VPC di Amazon MSK connessa a un client in un account diverso. AWS



Passaggio 1: sul cluster MSK nell'account A, attiva la connettività multi-VPC per lo schema di autenticazione IAM sul cluster

Il proprietario del cluster MSK deve configurare le impostazioni di configurazione sul cluster MSK dopo la creazione del cluster e in uno stato ACTIVE.

Il proprietario del cluster attiva la connettività privata multi-VPC sul cluster ACTIVE per tutti gli schemi di autenticazione che saranno attivi sul cluster. Questa operazione può essere eseguita utilizzando l'API o la console MSKUpdateSecurity . La connettività privata multi-VPC supporta gli schemi di autenticazione IAM, TLS e SASL/SCRAM. La connettività privata multi-VPC non può essere abilitata per i cluster non autenticati.

In questo caso d'uso, configurerai il cluster per utilizzare lo schema di autenticazione IAM.

> **Note**
>
> Se stai configurando il cluster MSK per utilizzare lo schema di autenticazione SASL/SCRAM, è obbligatorio utilizzare la proprietà "`allow.everyone.if.no.acl.found=false`" per le ACL di Apache Kafka. Consulta la pagina Apache Kafka ACLs.

Quando aggiorni le impostazioni di connettività privata multi-VPC, Amazon MSK intraprende un riavvio progressivo dei nodi del broker che aggiorna le configurazioni del broker. Il completamento del processo può richiedere fino a 30 minuti o più. Non è possibile apportare altri aggiornamenti al cluster durante l'aggiornamento della connettività.

Attivazione del multi-VPC per gli schemi di autenticazione selezionati sul cluster nell'account A tramite la console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/ per l'account in cui si trova il cluster.
2. Nel riquadro di navigazione, in Cluster MSK, scegli Cluster per visualizzare l'elenco dei cluster presenti nell'account.
3. Seleziona il cluster da configurare per la connettività privata multi-VPC. Il cluster deve essere in uno stato ACTIVE.
4. Seleziona la scheda Proprietà del cluster, quindi vai a Impostazioni di rete.
5. Seleziona il menu a discesa Modifica e seleziona Attiva la connettività multi-VPC.
6. Seleziona uno o più tipi di autenticazione che desideri attivare per questo cluster. Per questo caso d'uso, seleziona l'autenticazione basata sui ruoli IAM.
7. Seleziona Salva modifiche.

Example - UpdateConnectivity API che attiva schemi di autenticazione della connettività privata multi-VPC su un cluster

In alternativa alla console MSK, è possibile utilizzare l'UpdateConnectivity API per attivare la connettività privata multi-VPC e configurare gli schemi di autenticazione su un cluster ACTIVE. L'esempio seguente mostra lo schema di autenticazione IAM attivato per il cluster.

```
{
  "currentVersion": "K3T4TT2Z381HKD",
  "connectivityInfo": {
    "vpcConnectivity": {
      "clientAuthentication": {
        "sasl": {
          "iam": {
            "enabled": TRUE
            }
        }
      }
    }
  }
}
```

Amazon MSK crea l'infrastruttura di rete necessaria per la connettività privata. Amazon MSK crea anche un nuovo set di endpoint broker di bootstrap per ogni tipo di autenticazione che richiede la connettività privata. Tieni presente che lo schema di autenticazione non crittografata non supporta la connettività privata multi-VPC.

Passaggio 2: collegamento di una policy del cluster al cluster MSK

Il proprietario del cluster può collegare una policy del cluster (nota anche come policy basata sulle risorse) al cluster MSK in cui verrà attivata la connettività privata multi-VPC. La policy del cluster fornisce ai client l'autorizzazione ad accedere al cluster da un altro account. Prima di poter modificare la policy del cluster, sono necessari gli ID degli account che devono disporre dell'autorizzazione ad accedere al cluster MSK. Consulta la sezione Funzionamento di Amazon MSK con IAM.

Il proprietario del cluster deve collegare al cluster MSK una policy del cluster che autorizzi l'utente multi-account nell'account B a recuperare i broker di bootstrap per il cluster e ad autorizzare le seguenti operazioni sul cluster MSK nell'account A:

- CreateVpcConnessione
- GetBootstrapBroker
- DescribeCluster
- DescribeClusterV2

Example

A titolo di riferimento, di seguito è riportato un esempio di JSON per una policy di cluster di base, simile alla policy predefinita mostrata nell'editor di policy IAM della console MSK.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": [
                    "123456789012"
                ]
            },
            "Action": [
                "kafka:CreateVpcConnection",
                "kafka:GetBootstrapBrokers",
```

```
                "kafka:DescribeCluster",
                "kafka:DescribeClusterV2"
            ],
            "Resource": "arn:aws:kafka:us-east-1:123456789012:cluster/testing/
de8982fa-8222-4e87-8b20-9bf3cdfa1521-2"
        }
    ]
}
```

Collegamento di una policy del cluster al cluster MSK

1. Nella console Amazon MSK, in Cluster MSK, scegli Cluster.
2. Scorri verso il basso fino a Impostazioni di sicurezza e seleziona Modifica policy del cluster.
3. Nella console, nella schermata Modifica policy del cluster, seleziona Policy di base per la connettività multi-VPC.
4. Nel campo ID account, inserisci l'ID account per ogni account che dovrebbe disporre dell'autorizzazione per accedere a questo cluster. Durante la digitazione, l'ID viene automaticamente copiato nella sintassi JSON della policy visualizzata. Nel nostro esempio di policy del cluster, l'ID account è `123456789012`.
5. Seleziona Salva modifiche.

Per informazioni sulle API delle policy del cluster, consulta la sezione Policy basate sulle risorse di Amazon MSK.

Passaggio 3: operazioni dell'utente multi-account per configurare connessioni VPC gestite dal client

Per configurare la connettività privata multi-VPC tra un client in un account diverso dal cluster MSK, l'utente multi-account crea una connessione VPC gestita per il client. È possibile connettere più client al cluster MSK ripetendo questa procedura. Ai fini di questo caso d'uso, configurerai un solo client.

I client possono utilizzare gli schemi di autenticazione supportati IAM, SASL/SCRAM o TLS. A ogni connessione VPC gestita può essere associato un solo schema di autenticazione. Lo schema di autenticazione del client deve essere configurato nel cluster MSK a cui il client si connetterà.

In questo caso d'uso, configura lo schema di autenticazione del client in modo che il client nell'account B utilizzi lo schema di autenticazione IAM.

Prerequisiti

Questo processo richiede i seguenti elementi:

- La policy del cluster creata in precedenza che concede al client dell'account B l'autorizzazione a eseguire operazioni sul cluster MSK nell'account A.
- Una politica di identità allegata al client nell'Account B che concede autorizzazioni `kafka:CreateVpcConnection` e azioni. `ec2:CreateTags` `ec2:CreateVPCEndpoint` `ec2:DescribeVpcAttribute`

Example

A titolo di riferimento, di seguito è riportato un esempio di JSON per una policy di identità del client di base.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kafka:CreateVpcConnection",
        "ec2:CreateTags",
        "ec2:CreateVPCEndpoint",
        "ec2:DescribeVpcAttribute"
      ],
      "Resource": "*"
    }
  ]
}
```

Creazione di una connessione VPC gestita per un client nell'account B

1. Dall'amministratore del cluster, ottieni l'ARN del cluster MSK nell'account A al quale desideri che il client nell'account B si connetta. Prendi nota dell'ARN del cluster da utilizzare in seguito.
2. Nella console MSK per l'account client B, scegli Connessioni VPC gestite, quindi scegli Crea connessione.
3. Nel riquadro Impostazioni di connessione, incolla l'ARN del cluster nel campo di testo ARN del cluster, quindi scegli Verifica.
4. Seleziona Tipo di autenticazione per il client nell'account B. Per questo caso d'uso, scegli IAM quando crei la connessione VPC del client.
5. Scegli il VPC per il client.

6. Scegli almeno due zone di disponibilità e sottoreti associate. È possibile ottenere gli ID delle zone di disponibilità dai dettagli del cluster della console di AWS gestione o utilizzando l'DescribeClusterAPI o il comando AWS CLI describe-cluster. Gli ID di zona specificati per la sottorete client devono corrispondere a quelli della sottorete del cluster. Se mancano i valori per una sottorete, crea innanzitutto una sottorete con lo stesso ID di zona del cluster MSK.
7. Scegli un Gruppo di sicurezza per questa connessione VPC. È possibile accettare il gruppo di sicurezza predefinito. Per ulteriori informazioni sulla configurazione di un gruppo di sicurezza, consulta la pagina Control traffic to resources using security groups.
8. Seleziona Crea connessione.
9. Per ottenere l'elenco delle nuove stringhe del broker di bootstrap dalla console MSK dell'utente multi-account (Dettagli del cluster > Connessione VPC gestita), consulta le stringhe del broker di bootstrap mostrate in "Stringa di connessione al cluster". Dall'account B del cliente, l'elenco dei broker bootstrap può essere visualizzato chiamando l'API Brokers o visualizzando l'elenco GetBootstrapdei broker bootstrap nei dettagli del cluster della console.
10. Aggiorna i gruppi di sicurezza associati alle connessioni VPC come segue:

    a. Imposta le regole in entrata per il PrivateLink VPC per consentire tutto il traffico per l'intervallo IP dalla rete dell'Account B.

    b. [Facoltativo] Imposta la connettività delle Regole in uscita al cluster MSK. Scegli il Gruppo di sicurezza nella console VPC, Modifica le regole in uscita e aggiungi una regola per il Traffico TCP personalizzato per gli intervalli di porte 14001-14100. Il Network Load Balancer multi-VPC è in ascolto sugli intervalli di porte 14001-14100. Consulta la pagina Network Load Balancer.
11. Configura il client nell'account B per utilizzare i nuovi broker di bootstrap per la connettività privata multi-VPC per connettersi al cluster MSK nell'account A. Consulta la sezione Produzione e utilizzo di dati.

Una volta completata l'autorizzazione, Amazon MSK crea una connessione VPC gestita per ogni VPC e schema di autenticazione specificati. Il gruppo di sicurezza scelto è associato a ciascuna connessione. Questa connessione VPC gestita è configurata da Amazon MSK per connettersi privatamente ai broker. Puoi utilizzare il nuovo set di broker di bootstrap per connetterti privatamente al cluster Amazon MSK.

Aggiornamento degli schemi di autorizzazione su un cluster

La connettività privata multi-VPC supporta vari schemi di autenticazione: SASL/SCRAM, IAM, and TLS. Il proprietario del cluster può attivare/disattivare la connettività privata per uno o più schemi di autenticazione. Il cluster deve essere in stato ACTIVE per eseguire questa operazione.

Attivazione di uno schema di autenticazione tramite la console Amazon MSK

1. Apri la console Amazon MSK all'indirizzo AWS Management Console per il cluster che desideri modificare.
2. Nel riquadro di navigazione, in Cluster MSK, scegli Cluster per visualizzare l'elenco dei cluster presenti nell'account.
3. Seleziona il cluster da modificare. Il cluster deve essere in uno stato ACTIVE.
4. Seleziona la scheda Proprietà del cluster, quindi vai a Impostazioni di rete.
5. Seleziona il menu a discesa Modifica e seleziona Attiva la connettività multi-VPC per attivare un nuovo schema di autenticazione.
6. Seleziona uno o più tipi di autenticazione che desideri attivare per questo cluster.
7. Seleziona Attiva la selezione.

Quando attivi un nuovo schema di autenticazione, dovresti anche creare nuove connessioni VPC gestite per il nuovo schema di autenticazione e aggiornare i client di modo che utilizzino i broker di bootstrap specifici per il nuovo schema di autenticazione.

Disattivazione di uno schema di autenticazione tramite la console Amazon MSK

> **Note**
>
> Quando si disattiva la connettività privata multi-VPC per gli schemi di autenticazione, tutte le infrastrutture relative alla connettività, incluse le connessioni VPC gestite, vengono eliminate.

Quando si disattiva la connettività privata multi-VPC per gli schemi di autenticazione, le connessioni VPC esistenti sul lato client diventano INACTIVE e l'infrastruttura PrivateLink sul lato cluster, incluse le connessioni VPC gestite, viene rimossa. L'utente multi-account può eliminare solo la connessione VPC inattiva. Se sul cluster viene riattivata la connettività privata, l'utente multi-account deve creare una nuova connessione al cluster.

1. Apri la console Amazon MSK all'indirizzo AWS Management Console.
2. Nel riquadro di navigazione, in Cluster MSK, scegli Cluster per visualizzare l'elenco dei cluster presenti nell'account.
3. Seleziona il cluster da modificare. Il cluster deve essere in uno stato ACTIVE.
4. Seleziona la scheda Proprietà del cluster, quindi vai a Impostazioni di rete.
5. Seleziona il menu a discesa Modifica e seleziona Disattiva la connettività multi-VPC per disattivare uno schema di autenticazione.
6. Seleziona uno o più tipi di autenticazione che desideri disattivare per questo cluster.
7. Seleziona Disattiva la selezione.

Example Attivazione/disattivazione di uno schema di autenticazione tramite l'API

In alternativa alla console MSK, è possibile utilizzare l'UpdateConnectivity API per attivare la connettività privata multi-VPC e configurare gli schemi di autenticazione su un cluster ACTIVE. L'esempio seguente mostra gli schemi di autenticazione SASL/SCRAM e IAM attivati per il cluster.

Quando attivi un nuovo schema di autenticazione, dovresti anche creare nuove connessioni VPC gestite per il nuovo schema di autenticazione e aggiornare i client di modo che utilizzino i broker di bootstrap specifici per il nuovo schema di autenticazione.

Quando si disattiva la connettività privata multi-VPC per gli schemi di autenticazione, le connessioni VPC esistenti sul lato client diventano INACTIVE e l'infrastruttura PrivateLink sul lato cluster, incluse le connessioni VPC gestite, viene rimossa. L'utente multi-account può eliminare solo la connessione VPC inattiva. Se sul cluster viene riattivata la connettività privata, l'utente multi-account deve creare una nuova connessione al cluster.

```
Request:
{
  "currentVersion": "string",
  "connnectivityInfo": {
    "publicAccess": {
      "type": "string"
    },
    "vpcConnectivity": {
      "clientAuthentication": {
        "sasl": {
          "scram": {
```

```
            "enabled": TRUE
          },
          "iam": {
            "enabled": TRUE
          }
        },
        "tls": {
          "enabled": FALSE
        }
      }
    }
  }
}

Response:
{
  "clusterArn": "string",
  "clusterOperationArn": "string"
}
```

Rifiuto di una connessione VPC gestita a un cluster Amazon MSK

Dalla console Amazon MSK sull'account amministratore del cluster, puoi rifiutare una connessione VPC client. La connessione VPC del client deve essere nello stato AVAILABLE per essere rifiutata. Potresti voler rifiutare una connessione VPC gestita da un client che non è più autorizzato a connettersi al tuo cluster. Per evitare che nuove connessioni VPC gestite si connettano a un client, rifiuta l'accesso al client nella policy del cluster. Una connessione rifiutata comporta comunque dei costi fino a quando non viene eliminata dal proprietario della connessione. Consulta la sezione Eliminazione di una connessione VPC gestita a un cluster Amazon MSK.

Rifiuto di una connessione VPC client tramite la console MSK

1. Apri la console Amazon MSK all'indirizzo AWS Management Console.
2. Nel riquadro di navigazione, seleziona Cluster e scorri fino all'elenco Impostazioni di rete > Connessioni VPC client.
3. Seleziona la connessione che desideri rifiutare e seleziona Rifiuta connessione VPC client.
4. Conferma il rifiuto della connessione VPC client selezionata.

Per rifiutare una connessione VPC gestita tramite l'API, utilizza l'API `RejectClientVpcConnection`.

Eliminazione di una connessione VPC gestita a un cluster Amazon MSK

L'utente multi-account può eliminare una connessione VPC gestita per un cluster MSK dalla console dell'account client. Poiché l'utente proprietario del cluster non possiede la connessione VPC gestita, la connessione non può essere eliminata dall'account amministratore del cluster. Una volta eliminata, una connessione VPC non comporta più costi.

Eliminazione di una connessione VPC tramite la console MSK

1. Dall'account client, apri la console Amazon MSK all'indirizzo AWS Management Console.
2. Nel riquadro di navigazione, seleziona Connessioni VPC gestite.
3. Dall'elenco delle connessioni, seleziona la connessione VPN da eliminare.
4. Conferma l'eliminazione della connessione VPC.

Per eliminare una connessione VPC gestita tramite l'API, utilizza l'API `DeleteVpcConnection`.

Autorizzazioni per la connettività privata multi-VPC

Questa sezione riassume le autorizzazioni necessarie per client e cluster che utilizzano la funzionalità di connettività privata multi-VPC. La connettività privata multi-VPC richiede che l'amministratore del client crei le autorizzazioni su ogni client che avrà una connessione VPC gestita al cluster MSK. Richiede inoltre che l'amministratore del cluster MSK abiliti la PrivateLink connettività sul cluster MSK e selezioni gli schemi di autenticazione per controllare l'accesso al cluster.

Autenticazione del cluster e autorizzazioni di accesso all'argomento

Attiva la funzionalità di connettività privata multi-VPC per gli schemi di autenticazione abilitati per il tuo cluster MSK. Per informazioni, consulta Requisiti e limitazioni per la connettività privata multi-VPC. Se stai configurando il cluster MSK per utilizzare lo schema di autenticazione SASL/SCRAM, è obbligatorio utilizzare la proprietà `allow.everyone.if.no.acl.found=false` per le ACL di Apache Kafka. Dopo avere impostato le ACL Apache Kafka per il cluster, aggiorna la configurazione del cluster in modo che la proprietà `allow.everyone.if.no.acl.found` del cluster sia impostata su false. Per informazioni su come aggiornare la configurazione di un cluster, consulta la pagina Operazioni di configurazione di Amazon MSK.

Autorizzazioni delle policy del cluster multi-account

Se un client Kafka utilizza un AWS account diverso dal cluster MSK, allega al cluster MSK una policy basata sul cluster MSK che autorizzi l'utente root del client alla connettività tra account. È

possibile modificare la policy del cluster multi-VPC tramite l'editor di policy IAM nella console MSK (Impostazioni di sicurezza del cluster > Modifica policy del cluster) o utilizzare le seguenti API per gestire la policy del cluster:

PutClusterPolitica

Collega la policy del cluster al cluster. È possibile utilizzare questa API per creare o aggiornare la policy del cluster MSK specificata. Se stai aggiornando la policy, il campo CurrentVersion è obbligatorio nel payload della richiesta.

GetClusterPolitica

Recupera il testo JSON del documento di policy del cluster collegato al cluster.

DeleteClusterPolitica

Elimina la policy del cluster.

Di seguito è riportato un esempio di JSON per una policy di cluster di base, simile a quella mostrata nell'editor di policy IAM della console MSK.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": [
                    "123456789012"
                ]
            },
            "Action": [
                "kafka:CreateVpcConnection",
                "kafka:GetBootstrapBrokers",
                "kafka:DescribeCluster",
                "kafka:DescribeClusterV2"
            ],
            "Resource": "arn:aws:kafka:us-east-1:123456789012:cluster/testing/
de8982fa-8222-4e87-8b20-9bf3cdfa1521-2"
        }
    ]
}
```

Autorizzazioni client per la connettività privata multi-VPC a un cluster MSK

Per configurare la connettività privata multi-VPC tra un client Kafka e un cluster MSK, il client richiede una policy di identità collegata che conceda autorizzazioni per le operazioni `kafka:CreateVpcConnection`, `ec2:CreateTags` e `ec2:CreateVPCEndpoint` sul client. A titolo di riferimento, di seguito è riportato un esempio di JSON per una policy di identità del client di base.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kafka:CreateVpcConnection",
        "ec2:CreateTags",
        "ec2:CreateVPCEndpoint"
      ],
      "Resource": "*"
    }
  ]
}
```

Informazioni sulle porte

Utilizza i seguenti numeri di porta in modo che Amazon MSK possa comunicare con i computer client:

- Per comunicare con i broker con testo non crittografato, utilizza la porta 9092.
- Per comunicare con i broker con crittografia TLS, utilizzate la porta 9094 per l'accesso dall'interno AWS e la porta 9194 per l'accesso pubblico.
- Per comunicare con i broker con SASL/SCRAM, utilizzate la porta 9096 per l'accesso dall'interno e la porta 9196 per l'accesso pubblico. AWS
- Per comunicare con i broker in un cluster configurato per l'usothe section called “Controllo degli accessi IAM”, utilizzate la porta 9098 per l'accesso dall'interno e la porta 9198 per l'accesso pubblico. AWS
- Per comunicare con Apache ZooKeeper utilizzando la crittografia TLS, utilizzate la porta 2182. ZooKeeper I nodi Apache utilizzano la porta 2181 per impostazione predefinita.

Migrazione a un cluster Amazon MSK

Il replicatore Amazon MSK può essere utilizzato per eseguire la migrazione dei cluster MSK. Per informazioni, consulta Cos'è il replicatore Amazon MSK?. In alternativa, puoi utilizzare Apache MirrorMaker 2.0 per migrare da un cluster non MSK a un cluster Amazon MSK. Per un esempio di come eseguire questa operazione, consulta Migrare un cluster Apache Kafka locale su Amazon MSK utilizzando. MirrorMaker Per informazioni sull'uso MirrorMaker, consulta Mirroring dei dati tra i cluster nella documentazione di Apache Kafka. Ti consigliamo di eseguire la configurazione in una configurazione ad alta MirrorMaker disponibilità.

Una descrizione dei passaggi da seguire quando si utilizza per MirrorMaker migrare a un cluster MSK

1. Creazione del cluster MSK di destinazione
2. Inizia MirrorMaker da un'istanza Amazon EC2 all'interno dello stesso Amazon VPC del cluster di destinazione.
3. Ispeziona il ritardo. MirrorMaker
4. Dopo aver MirrorMaker recuperato il ritardo, reindirizza produttori e consumatori al nuovo cluster utilizzando i broker bootstrap del cluster MSK.
5. MirrorMakerSpegnere.

Migrazione del cluster Apache Kafka ad Amazon MSK

Supponi di disporre di un cluster Apache Kafka denominato `CLUSTER_ONPREM`, popolato con argomenti e dati. Se desideri eseguire la migrazione di tale cluster in un nuovo cluster Amazon MSK denominato `CLUSTER_AWSMSK`, questa procedura fornisce una vista generale dei passaggi necessari.

Migrazione del cluster Apache Kafka esistente ad Amazon MSK

1. In `CLUSTER_AWSMSK`, creare tutti gli argomenti che desideri migrare.

   Non puoi utilizzarlo MirrorMaker per questo passaggio perché non ricrea automaticamente gli argomenti che desideri migrare con il giusto livello di replica. È possibile creare gli argomenti in Amazon MSK con gli stessi fattori di replica e numeri di partizioni che esistevano in `CLUSTER_ONPREM`. È possibile inoltre creare gli argomenti con diversi fattori di replica e numeri di partizioni.

2. Inizia MirrorMaker da un'istanza con accesso in lettura `CLUSTER_ONPREM` e accesso in scrittura. `CLUSTER_AWSMSK`
3. Eseguire il comando seguente per creare una copia speculare di tutti gli argomenti:

   ```
   <path-to-your-kafka-installation>/bin/kafka-mirror-maker.sh --consumer.config
    config/mirrormaker-consumer.properties --producer.config config/mirrormaker-
   producer.properties --whitelist '.*'
   ```

   In questo comando, `config/mirrormaker-consumer.properties` punta a un broker bootstrap in `CLUSTER_ONPREM`; ad esempio, `bootstrap.servers=localhost:9092`. E `config/mirrormaker-producer.properties` indica un broker di bootstrap in CLUSTER_AWSMSK; ad esempio,. `bootstrap.servers=10.0.0.237:9092,10.0.2.196:9092,10.0.1.233:9092`
4. Continua a MirrorMaker funzionare in background e continua a utilizzare. `CLUSTER_ONPREM` MirrorMaker rispecchia tutti i nuovi dati.
5. Controlla lo stato di avanzamento del mirroring controllando il ritardo tra l'ultimo offset di ogni argomento e l'offset corrente da cui si sta consumando. MirrorMaker

   Ricorda che si MirrorMaker tratta semplicemente di utilizzare un consumatore e un produttore. Quindi, è possibile controllare il ritardo usando lo strumento `kafka-consumer-groups.sh`. Per trovare il nome del gruppo di consumatori, cercare all'interno del file `group.id` `mirrormaker-consumer.properties` e utilizzare il suo valore. Se tale chiave non esiste nel file, è possibile crearla. Ad esempio, impostare `group.id=mirrormaker-consumer-group`.
6. Dopo aver MirrorMaker finito di rispecchiare tutti gli argomenti, interrompete tutti i produttori e i consumatori, e poi smettete MirrorMaker. Quindi, reindirizzare i produttori e i consumatori al cluster `CLUSTER_AWSMSK` modificando i relativi valori dei broker bootstrap del produttore e del consumatore. Riavviare tutti i produttori e i consumatori su `CLUSTER_AWSMSK`.

Migrazione da un cluster Amazon MSK a un altro

È possibile utilizzare Apache MirrorMaker 2.0 per migrare da un cluster non MSK a un cluster MSK. Ad esempio, puoi eseguire la migrazione da una versione di Apache Kafka a un'altra. Per un esempio di come eseguire questa operazione, consulta Migrare un cluster Apache Kafka locale su Amazon MSK utilizzando. MirrorMaker In alternativa, è possibile utilizzare il replicatore Amazon MSK per eseguire la migrazione dei cluster MSK. Per ulteriori informazioni sul replicatore Amazon MSK, consulta la pagina Replicatore MSK.

MirrorMaker 1.0 migliori pratiche

Questo elenco di best practice si applica alla MirrorMaker versione 1.0.

- Esegui MirrorMaker sul cluster di destinazione. In questo modo, se si verifica un problema di rete, i messaggi sono ancora disponibili nel cluster di origine. Se esegui MirrorMaker sul cluster di origine e gli eventi sono memorizzati nel buffer nel produttore e c'è un problema di rete, gli eventi potrebbero andare persi.
- Se è richiesta la crittografia dei dati in transito, eseguirla nel cluster di origine.
- Per i consumatori, impostare auto.commit.enabled=false
- Per i produttori, impostare
  - max.in.flight.requests.per.connection=1
  - retries=Int.Max_Value
  - acks=all
  - max.block.ms = Long.Max_Value
- Per un elevato throughput produttore:
  - Esegui il buffer di messaggi e compila batch di messaggi: tune buffer.memory, batch.size, linger.ms
  - Ottimizza i buffer dei socket: receive.buffer.bytes, send.buffer.bytes
- Per evitare la perdita di dati, disattiva il commit automatico all'origine, in modo che MirrorMaker possa controllare i commit, cosa che in genere esegue dopo aver ricevuto l'ack dal cluster di destinazione. Se il produttore ha acks=all e il cluster di destinazione ha impostato min.insync.replicas su più di 1, i messaggi vengono mantenuti su più di un broker nella destinazione prima che il consumatore esegua il commit dell'offset all'origine. MirrorMaker
- Se l'ordine è importante, puoi impostare i tentativi su 0. In alternativa, per un ambiente di produzione, imposta il numero massimo di connessioni in transito su 1 per garantire che non venga eseguito il commit dei batch inviati senza seguire un ordine se un batch non riesce a metà. In questo modo, ogni batch inviato viene ritentato finché il batch successivo non viene inviato. Se max.block.ms non è impostato sul valore massimo e se il buffer del produttore è pieno, potrebbe verificarsi una perdita di dati (a seconda di alcune delle altre impostazioni). Questo può bloccare e causare uno stato di congestione nel consumatore.
- Per elevato throughput
  - Incrementa buffer.memory.
  - Incrementa le dimensioni batch.

  - Ottimizza linger.ms per consentire il riempimento dei batch. Ciò consente inoltre una migliore compressione, meno utilizzo della larghezza di banda della rete e meno storage sul cluster. Questo comporta un aumento della conservazione.
  - Monitora l'utilizzo della CPU e della memoria.
- Per elevato throughput consumatore
  - Aumenta MirrorMaker il numero di thread/consumatori per processo: num.streams.
  - Aumenta il numero di MirrorMaker processi tra le macchine prima di aumentare i thread per consentire un'elevata disponibilità.
  - Aumenta il numero di MirrorMaker processi prima sulla stessa macchina e poi su macchine diverse (con lo stesso ID di gruppo).
  - Isola gli argomenti con una velocità effettiva molto elevata e utilizza istanze separate MirrorMaker .
- Per gestione e configurazione
  - Strumenti di gestione AWS CloudFormation dell'uso e della configurazione come Chef e Ansible.
  - Utilizza montaggi Amazon EFS per mantenere tutti i file di configurazione accessibili da tutte le istanze Amazon EC2.
  - Utilizza i contenitori per semplificare la scalabilità e la gestione delle MirrorMaker istanze.
- In genere, è necessario più di un consumatore per saturare un produttore. MirrorMaker Pertanto, configura più consumatori. Innanzitutto, configurali su macchine diverse per fornire elevata disponibilità. Quindi, dimensiona le singole macchine fino ad avere un consumatore per ogni partizione, con i consumatori distribuiti in modo uniforme tra le macchine.
- Per elevato throughput di inserimento e consegna, ottimizza i buffer di ricezione e invio perché le relative impostazione predefinite potrebbero essere troppo basse. Per ottenere le massime prestazioni, assicuratevi che il numero totale di stream (num.streams) corrisponda a tutte le partizioni degli argomenti che state tentando di copiare nel MirrorMaker cluster di destinazione.

MirrorMaker 2.* vantaggi

- Utilizza il framework e l'ecosistema Apache Kafka Connect.
- Rileva nuovi argomenti e partizioni.
- Sincronizza automaticamente la configurazione degli argomenti tra cluster.
- Supporta coppie di cluster "attiva/attiva", così come qualsiasi numero di cluster attivi.
- Fornisce nuove metriche, tra cui end-to-end la latenza di replica su più data center e cluster.

- Emette gli offset necessari per eseguire la migrazione dei consumatori tra cluster e fornisce strumenti per la traslazione dell'offset.
- Supporta un file di configurazione di alto livello per specificare più cluster e flussi di replica in un'unica posizione, rispetto alle proprietà produttore/consumatore di basso livello per ogni processo 1.*. MirrorMaker

Monitoraggio di un cluster Amazon MSK

Esistono diversi modi in cui Amazon MSK consente di monitorare lo stato del cluster Amazon MSK.

- Amazon MSK ti aiuta a monitorare la capacità di archiviazione su disco, inviando automaticamente avvisi quando un cluster sta per raggiungere il limite di capacità di archiviazione. Gli avvisi forniscono anche raccomandazioni sulle misure migliori da intraprendere per risolvere i problemi rilevati. Ciò consente di identificare e risolvere rapidamente i problemi relativi alla capacità del disco prima che diventino critici. Amazon MSK invia automaticamente questi avvisi alla console Amazon MSK, ad AWS Health Dashboard Amazon EventBridge e ai contatti e-mail del tuo account. AWS Per ulteriori informazioni sull'aumento della capacità di archiviazione, consulta Avvisi sulla capacità di archiviazione di Amazon MSK.
- Amazon MSK raccoglie i parametri di Apache Kafka e li invia ad Amazon CloudWatch dove puoi visualizzarli. Per ulteriori informazioni sui parametri Apache Kafka, inclusi quelli esposti da Amazon MSK, consulta la pagina Monitoring nella documentazione di Apache Kafka.
- Puoi anche monitorare il cluster MSK con Prometheus, un'applicazione di monitoraggio open source. Per informazioni su Prometheus, consulta la sezione relativa alla panoramica nella documentazione di Prometheus. Per informazioni su come monitorare il cluster con Prometheus, consulta the section called “Monitoraggio aperto con Prometheus”.

Argomenti

- Metriche di Amazon MSK per il monitoraggio con CloudWatch
- Visualizzazione dei parametri di Amazon MSK utilizzando CloudWatch
- Monitoraggio del ritardo dei consumatori
- Monitoraggio aperto con Prometheus
- Avvisi sulla capacità di archiviazione di Amazon MSK

Metriche di Amazon MSK per il monitoraggio con CloudWatch

Amazon MSK si integra con Amazon per CloudWatch consentirti di raccogliere, visualizzare e analizzare i CloudWatch parametri per il tuo cluster Amazon MSK. Le metriche configurate per il cluster MSK vengono raccolte e inviate automaticamente. CloudWatch Puoi impostare il livello di monitoraggio per un cluster MSK su uno dei seguenti valori: `DEFAULT`, `PER_BROKER`,

PER_TOPIC_PER_BROKER o PER_TOPIC_PER_PARTITION. Le tabelle nelle sezioni seguenti mostrano tutti i parametri resi disponibili a partire da ciascun livello di monitoraggio.

> **Note**
>
> I nomi di alcuni parametri di Amazon MSK per il CloudWatch monitoraggio sono cambiati nella versione 3.6.0 e successive. Usa i nuovi nomi per monitorare questi parametri. Per i parametri con nomi modificati, la tabella seguente mostra il nome utilizzato nella versione 3.6.0 e successive, seguito dal nome nella versione 2.8.2.tiered.

I parametri del livello DEFAULT sono gratuiti. I prezzi per altre metriche sono descritti nella pagina CloudWatchdei prezzi di Amazon.

Monitoraggio del livello **DEFAULT**

I parametri descritti nella tabella seguente sono disponibili a livello di monitoraggio DEFAULT e sono gratuiti.

Parametri disponibili al livello di monitoraggio **DEFAULT**

Nome	Quando visibile	Dimensi i	Descrizione
ActiveController Count	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome del cluster	Solo un controller per cluster deve essere attivo in qualsiasi momento.
BurstBalance	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il saldo residuo dei crediti di espansione input-output per i volumi EBS nel cluster. Utilizzalo per analizzare la latenza o la riduzione della velocità di trasmissione effettiva. BurstBalance non viene riportato per i volumi EBS quando le prestazioni di base di un volume sono maggiori delle prestazioni massime di espansione. Per ulteriori informazioni,

Nome	Quando visibile	Dimensi i	Descrizione
			consulta la pagina I/O Credits and burst performance.
`BytesInPerSec`	Dopo aver creato un argomento.	Nome cluster, ID broker, argomen	Il numero di byte al secondo ricevuti dai client. Questo parametro è disponibile per broker e anche per argomento.
`BytesOutPerSec`	Dopo aver creato un argomento.	Nome cluster, ID broker, argomen	Il numero di byte al secondo inviati ai client. Questo parametro è disponibile per broker e anche per argomento.
`ClientConnectionCount`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome del cluster, ID broker, autenticazione client	Il numero di connessioni client autenticate attive.
`ConnectionCount`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di connessioni attive autenticate, non autenticate e tra broker.

Nome	Quando visibile	Dimensioni	Descrizione
`CPUCreditBalance`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di crediti CPU ottenuti da un broker da quando è stato lanciato. I crediti vengono accumulati nel saldo del credito dopo che sono stati ottenuti e rimossi dal saldo del credito una volta spesi. L'esaurimento del credito della CPU può avere un impatto negativo sulle prestazioni del cluster. È possibile adottare delle misure per ridurre il carico della CPU. Ad esempio, puoi ridurre il numero di richieste dei client o aggiornare il tipo di broker a un tipo di broker M5.
`CpuIdle`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La percentuale di tempo di inattività della CPU.
`CpuIoWait`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La percentuale di inattività della CPU durante un'operazione su disco in sospeso.
`CpuSystem`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La percentuale di CPU nello spazio del kernel.
`CpuUser`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La percentuale di CPU nello spazio utente.

Nome	Quando visibile	Dimensi i	Descrizione
`GlobalPartitionCount`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome del cluster	Il numero di partizioni in tutti gli argomenti del cluster, escluse le repliche. Poiché `GlobalPartitionCount` non include le repliche, la somma dei `PartitionCount` valori può essere superiore a quella che si otterrebbe GlobalPartitionCount se il fattore di replica per un argomento fosse maggiore di 1.
`GlobalTopicCount`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome del cluster	Numero totale di argomenti in tutti i broker nel cluster.
`EstimatedMaxTimeLag`	Dopo che un gruppo di consumatori ha utilizzato un argomento.	Gruppo di consuma ri, argomen	Tempo stimato (in secondi) per lo svuotamento di `MaxOffsetLag` .
`KafkaAppLogsDiskUsed`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La percentuale di spazio su disco utilizzata per i log delle applicazioni.
`KafkaDataLogsDiskUsed` (dimensione `Cluster Name, Broker ID`)	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La percentuale di spazio su disco utilizzato per i log dei dati.
`KafkaDataLogsDiskUsed` (dimensione `Cluster Name`)	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome del cluster	La percentuale di spazio su disco utilizzato per i log dei dati.

Nome	Quando visibile	Dimensi i	Descrizione
`LeaderCount`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero totale di leader delle partizioni per broker, escluse le repliche.
`MaxOffsetLag`	Dopo che un gruppo di consumatori ha utilizzato un argomento.	Gruppo di consuma ri, argomen	Il ritardo massimo di offset su tutte le partizioni di un argomento.
`MemoryBuffered`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La dimensione in byte di memoria nel buffer per il broker.
`MemoryCached`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La dimensione in byte di memoria nella cache per il broker.
`MemoryFree`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La dimensione in byte di memoria libera e disponibile per il broker.
`HeapMemor yAfterGC`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La percentuale di memoria heap totale in uso dopo la rimozione di oggetti inutili.
`MemoryUsed`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La dimensione in byte di memoria utilizzata per il broker.

Nome	Quando visibile	Dimensioni	Descrizione
`MessagesInPerSec`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di messaggi in entrata al secondo per il broker.
`NetworkRxDropped`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di pacchetti ricezione eliminati.
`NetworkRxErrors`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di errori ricezione di rete per il broker.
`NetworkRxPackets`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di pacchetti ricevuti dal broker.
`NetworkTxDropped`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di pacchetti trasmissione eliminati.
`NetworkTxErrors`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di errori trasmissione di rete per il broker.
`NetworkTxPackets`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di pacchetti trasmessi dal broker.

Nome	Quando visibile	Dimensi i	Descrizione
`OfflinePartitionsCount`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome del cluster	Numero totale di partizioni che sono offline nel cluster.
`PartitionCount`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero totale di partizioni di argomento per broker, incluse le repliche.
`ProduceTotalTimeMsMean`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il tempo di produzione medio in millisecondi.
`RequestBytesMean`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero medio di byte della richiesta per il broker.
`RequestTime`	Dopo l'applicazione del throttling della richiesta .	Nome cluster, ID broker	Il tempo medio in millisecondi trascorso nella rete di broker e nei thread I/O per elaborare le richieste.
`RootDiskUsed`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La percentuale del disco radice utilizzato dal broker.
`SumOffsetLag`	Dopo che un gruppo di consumatori ha utilizzato un argomento.	Gruppo di consuma ri, argomen	Il ritardo di offset aggregato per tutte le partizioni di un argomento.

Nome	Quando visibile	Dimensi i	Descrizione
SwapFree	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La dimensione in byte della memoria swap disponibile per il broker.
SwapUsed	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	La dimensione in byte della memoria swap utilizzata dal broker.
TrafficShaping	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Parametri di alto livello che indicano il numero di pacchetti formati (abbandonati o messi in coda) a causa di un eccesso di allocazioni di rete. Maggiori dettagli sono disponibili con i parametri PER_BROKER.
UnderMinIsrPartitionCount	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di partizioni minIsr under per il broker.
UnderReplicatedPartitions	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Il numero di partizioni replicate per il broker.
ZooKeeperRequestLatencyMsMean	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Per cluster ZooKeeper basato. La latenza media in millisecondi per le richieste ZooKeeper Apache al broker.

Nome	Quando visibile	Dimensi i	Descrizione
`ZooKeeper SessionState`	Dopo che il cluster raggiunge lo stato ACTIVE.	Nome cluster, ID broker	Per cluster basato. ZooKeeper Stato della connessione della ZooKeeper sessione del broker che può essere una delle seguenti: NOT_CONNE CTED: '0.0', ASSOCIATING: '0.1', CONNECTING: '0.5', CONNECTED READONLY: '0.8', CONNECTED: '1.0', CLOSED: '5.0', AUTH_FAILED: '10.0'.

Monitoraggio del livello **PER_BROKER**

Quando imposti il livello di monitoraggio su `PER_BROKER`, ottieni i parametri descritti nella tabella seguente oltre a tutti i parametri del livello `DEFAULT`. Paghi per i parametri nella tabella seguente, mentre i parametri del livello `DEFAULT` continuano a essere gratuiti. I parametri contenuti in questa tabella hanno le seguenti dimensioni: Nome cluster, ID broker.

Parametri aggiuntivi disponibili a partire dal livello di monitoraggio **`PER_BROKER`**

Nome	Quando visibile	Descrizione
`BwInAllowanceExceeded`	Dopo che il cluster raggiunge lo stato ACTIVE.	Numero di pacchetti modellati perché la larghezza di banda aggregata in entrata ha superato il valore massimo per l'istanza.
`BwOutAllowanceExce eded`	Dopo che il cluster raggiunge lo stato ACTIVE.	Numero di pacchetti modellati perché la larghezza di banda aggregata in uscita ha superato il valore massimo per l'istanza.
`ConnTrackAllowance Exceeded`	Dopo che il cluster raggiunge lo stato ACTIVE.	Numero di pacchetti modellati perché il tracciamento della connessione ha superato il valore massimo per il broker. Il tracciamento della connessio

Nome	Quando visibile	Descrizione
		ne è legato ai gruppi di sicurezza che tengono traccia di ogni connessione stabilita per garantire che i pacchetti restituiti vengano consegnati come previsto.
`ConnectionCloseRate`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di connessioni chiuse al secondo per ascoltatore. Questo numero viene aggregato per ascoltatore e filtrato per gli ascoltatori client.
`ConnectionCreationRate`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di nuove connessioni stabilite al secondo per ascoltatore. Questo numero viene aggregato per ascoltatore e filtrato per gli ascoltatori client.
`CpuCreditUsage`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di crediti CPU utilizzati dal broker. L'esaurimento del credito della CPU può avere un impatto negativo sulle prestazioni del cluster. È possibile adottare delle misure per ridurre il carico della CPU. Ad esempio, puoi ridurre il numero di richieste dei client o aggiornare il tipo di broker a un tipo di broker M5.
`FetchConsumerLocalTimeMsMean`	Dopo che c'è un produttore/consumatore.	Tempo medio in millisecondi di elaborazione della richiesta del consumatore presso il leader.
`FetchConsumerRequestQueueTimeMsMean`	Dopo che c'è un produttore/consumatore.	Tempo medio in millisecondi di attesa della richiesta del consumatore nella coda delle richieste.

Nome	Quando visibile	Descrizione
FetchConsumerResponseQueueTimeMsMean	Dopo che c'è un produttore/consumatore.	Tempo medio in millisecondi di attesa della richiesta del consumatore nella coda delle risposte.
FetchConsumerResponseSendTimeMsMean	Dopo che c'è un produttore/consumatore.	Tempo medio in millisecondi impiegato dal consumatore per inviare una risposta.
FetchConsumerTotalTimeMsMean	Dopo che c'è un produttore/consumatore.	Il tempo totale medio in millisecondi impiegato dai consumatori per recuperare i dati dal broker.
FetchFollowerLocalTimeMsMean	Dopo che c'è un produttore/consumatore.	Tempo medio in millisecondi impiegato a livello di leader per elaborare la richiesta follower.
FetchFollowerRequestQueueTimeMsMean	Dopo che c'è un produttore/consumatore.	Tempo medio in millisecondi di attesa della richiesta follower nella coda delle richieste.
FetchFollowerResponseQueueTimeMsMean	Dopo che c'è un produttore/consumatore.	Tempo medio in millisecondi di attesa della richiesta follower nella coda delle risposte.
FetchFollowerResponseSendTimeMsMean	Dopo che c'è un produttore/consumatore.	Tempo medio in millisecondi impiegato dal follower per inviare una risposta.
FetchFollowerTotalTimeMsMean	Dopo che c'è un produttore/consumatore.	Il tempo totale medio in millisecondi impiegato dai follower per recuperare i dati dal broker.
FetchMessageConversionsPerSec	Dopo aver creato un argomento.	Il numero di conversioni dei messaggi di recupero al secondo per il broker.

Nome	Quando visibile	Descrizione
FetchThrottleByteRate	Dopo l'applicazione del throttling della larghezza di banda.	Il numero di byte sottoposti a throttling al secondo.
FetchThrottleQueueSize	Dopo l'applicazione del throttling della larghezza di banda.	Il numero di messaggi nella coda di throttling.
FetchThrottleTime	Dopo l'applicazione del throttling della larghezza di banda.	Il tempo medio del throttling di recupero in millisecondi.
IAMNumberOfConnectionRequests	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di richieste di autenticazione IAM al secondo.
IAMTooManyConnections	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di connessioni tentate è superiore a 100. 0 indica che il numero di connessioni rientra nel limite. Se >0, il limite di accelerazione viene superato ed è necessario ridurre il numero di connessioni.
NetworkProcessorAvgIdlePercent	Dopo che il cluster raggiunge lo stato ACTIVE.	La percentuale media del tempo di inattività dei processori di rete.
PpsAllowanceExceeded	Dopo che il cluster raggiunge lo stato ACTIVE.	Numero di pacchetti modellati perché il PPS bidirezionale ha superato il valore massimo per il broker.
ProduceLocalTimeMsMean	Dopo che il cluster raggiunge lo stato ACTIVE.	Tempo medio in millisecondi impiegato a livello di leader per elaborare la richiesta.
ProduceMessageConversionsPerSec	Dopo aver creato un argomento.	Il numero di conversioni di messaggi di produzione al secondo per il broker.

Nome	Quando visibile	Descrizione
ProduceMessageConversionsTimeMsMean	Dopo che il cluster raggiunge lo stato ACTIVE.	Il tempo medio in millisecondi impiegato per le conversioni di formato dei messaggi.
ProduceRequestQueueTimeMsMean	Dopo che il cluster raggiunge lo stato ACTIVE.	Il tempo medio in millisecondi che i messaggi di richiesta rimangono nella coda.
ProduceResponseQueueTimeMsMean	Dopo che il cluster raggiunge lo stato ACTIVE.	Il tempo medio in millisecondi che messaggi di risposta rimangono nella coda.
ProduceResponseSendTimeMsMean	Dopo che il cluster raggiunge lo stato ACTIVE.	Il tempo medio in millisecondi impiegato per l'invio di messaggi di risposta.
ProduceThrottleByteRate	Dopo l'applicazione del throttling della larghezza di banda.	Il numero di byte sottoposti a throttling al secondo.
ProduceThrottleQueueSize	Dopo l'applicazione del throttling della larghezza di banda.	Il numero di messaggi nella coda di throttling.
ProduceThrottleTime	Dopo l'applicazione del throttling della larghezza di banda.	Il tempo di throttling di produzione medio in millisecondi.
ProduceTotalTimeMsMean	Dopo che il cluster raggiunge lo stato ACTIVE.	Il tempo di produzione medio in millisecondi.

Nome	Quando visibile	Descrizione
`RemoteFetchBytesPerSec (RemoteBytesInPerSec in v2.8.2.tiered)`	Dopo che è presente un produttore/consumatore.	Il numero totale di byte trasferiti dall'archiviazione a più livelli in risposta alle richieste dei consumatori. Questo parametro include tutte le partizioni di argomento che contribuiscono al traffico di trasferimento dati a valle. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.
`RemoteCopyBytesPerSec (RemoteBytesOutPerSec in v2.8.2.tiered)`	Dopo che è presente un produttore/consumatore.	Il numero totale di byte trasferiti nell'archiviazione a più livelli, inclusi i dati provenienti da segmenti di log, indici e altri file ausiliari. Questo parametro include tutte le partizioni di argomento che contribuiscono al traffico di trasferimento dati a monte. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.
`RemoteLogManagerTasksAvgIdlePercent`	Dopo che il cluster raggiunge lo stato ACTIVE.	La percentuale media di tempo che il gestore di log remoto ha trascorso inattivo. Il gestore remoto dei log trasferisce i dati dal broker all'archiviazione a più livelli. Categoria: attività interna. Questo è un parametro KIP-405.
`RemoteLogReaderAvgIdlePercent`	Dopo che il cluster raggiunge lo stato ACTIVE.	La percentuale media di tempo che il lettore di log remoto ha trascorso inattivo. Il lettore di log remoto trasferisce i dati dall'archiviazione remota al broker in risposta alle richieste dei consumatori. Categoria: attività interna. Questo è un parametro KIP-405.

Nome	Quando visibile	Descrizione
`RemoteLogReaderTaskQueueSize`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di attività responsabili delle letture dall'archiviazione a più livelli in attesa di essere pianificate. Categoria: attività interna. Questo è un parametro KIP-405.
`RemoteFetchErrorsPerSec (RemoteReadErrorPerSec in v2.8.2.tiered)`	Dopo che il cluster raggiunge lo stato ACTIVE.	La percentuale totale di errori in risposta alle richieste di lettura che il broker specificato ha inviato all'archiviazione a più livelli per recuperare i dati in risposta alle richieste dei consumatori. Questo parametro include tutte le partizioni di argomento che contribuiscono al traffico di trasferimento dati a valle. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.
`RemoteFetchRequestsPerSec (RemoteReadRequestsPerSec in v2.8.2.tiered)`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero totale di richieste di lettura che il broker specificato ha inviato all'archiviazione a più livelli per recuperare i dati in risposta alle richieste dei consumatori. Questo parametro include tutte le partizioni di argomento che contribuiscono al traffico di trasferimento dati a valle. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.

Nome	Quando visibile	Descrizione
`RemoteCopyErrorsPerSec (RemoteWriteErrorPerSec in v2.8.2.tiered)`	Dopo che il cluster raggiunge lo stato ACTIVE.	La percentuale totale di errori in risposta alle richieste di scrittura che il broker specificato ha inviato all'archiviazione a più livelli per trasferire i dati a monte. Questo parametro include tutte le partizioni di argomento che contribuiscono al traffico di trasferimento dati a monte. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.
`ReplicationBytesInPerSec`	Dopo aver creato un argomento.	Il numero di byte al secondo ricevuti da altri broker.
`ReplicationBytesOutPerSec`	Dopo aver creato un argomento.	Il numero di byte al secondo inviati ad altri broker.
`RequestExemptFromThrottleTime`	Dopo l'applicazione del throttling della richiesta.	Il tempo medio in millisecondi trascorso nella rete di broker e nei thread I/O per elaborare le richieste esenti da throttling.
`RequestHandlerAvgIdlePercent`	Dopo che il cluster raggiunge lo stato ACTIVE.	La percentuale media del tempo di inattività dei thread del gestore di richieste.
`RequestThrottleQueueSize`	Dopo l'applicazione del throttling della richiesta.	Il numero di messaggi nella coda di throttling.
`RequestThrottleTime`	Dopo l'applicazione del throttling della richiesta.	Il tempo di throttling della richiesta medio in millisecondi.

Nome	Quando visibile	Descrizione
`TcpConnections`	Dopo che il cluster raggiunge lo stato ACTIVE.	Mostra il numero di segmenti TCP in entrata e in uscita con il flag SYN impostato.
`RemoteCopyLagBytes (TotalTierBytesLag in v2.8.2.tiered)`	Dopo aver creato un argomento.	Il numero totale di byte dei dati idonei per l'archiviazione a più livelli sul broker ma che non sono ancora stati trasferiti in tale archiviazione. Questi parametri mostrano l'efficienza del trasferimento dati a monte. Con l'aumentare del ritardo, aumenta la quantità di dati che non persistono nell'archiviazione a più livelli. Categoria : ritardo di archiviazione. Questo non è un parametro KIP-405.
`TrafficBytes`	Dopo che il cluster raggiunge lo stato ACTIVE.	Mostra il traffico di rete in byte complessivi tra client (produttori e consumatori) e broker. Il traffico tra i broker non viene segnalato.
`VolumeQueueLength`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di richieste di operazion i di lettura e scrittura in attesa di completamento nel periodo di tempo specificato.
`VolumeReadBytes`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di byte letti durante il periodo di tempo specificato.
`VolumeReadOps`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero totale di operazioni di lettura nel periodo di tempo specificato.

Nome	Quando visibile	Descrizione
`VolumeTotalReadTime`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero totale di secondi impiegato da tutte le operazioni di lettura completate nel periodo di tempo specificato.
`VolumeTotalWriteTime`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero totale di secondi impiegato da tutte le operazioni di scrittura completate nel periodo di tempo specificato.
`VolumeWriteBytes`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero di byte scritti durante il periodo di tempo specificato.
`VolumeWriteOps`	Dopo che il cluster raggiunge lo stato ACTIVE.	Il numero totale di operazioni di scrittura durante il periodo di tempo specificato.

Monitoraggio del livello **PER_TOPIC_PER_BROKER**

Quando imposti il livello di monitoraggio su `PER_TOPIC_PER_BROKER`, ottieni i parametri descritti nella tabella seguente, oltre a tutti i parametri dei livelli `PER_BROKER` e DEFAULT. Solo i parametri del livello `DEFAULT` sono gratuiti. I parametri contenuti in questa tabella hanno le seguenti dimensioni: Nome cluster, ID broker, Argomento.

> **Important**
>
> Per un cluster Amazon MSK che utilizza Apache Kafka 2.4.1 o una versione più recente, i parametri nella tabella seguente vengono visualizzati solo dopo che i loro valori diventano diversi da zero per la prima volta. Ad esempio, per visualizzare `BytesInPerSec`, uno o più produttori devono prima inviare i dati al cluster.

Parametri aggiuntivi disponibili a partire dal livello di monitoraggio **PER_TOPIC_PER_BROKER**

Nome	Quando visibile	Descrizione
FetchMessageConversionsPerSec	Dopo aver creato un argomento.	Il numero di messaggi recuperati convertiti al secondo.
MessagesInPerSec	Dopo aver creato un argomento.	Il numero di messaggi ricevuti al secondo.
ProduceMessageConversionsPerSec	Dopo aver creato un argomento.	Il numero di conversioni al secondo per i messaggi prodotti.
RemoteFetchBytesPerSec (RemoteBytesInPerSec in v2.8.2.tiered)	Dopo aver creato un argomento, l'argomento è in fase di produzione/utilizzo.	Il numero di byte trasferiti nell'archiviazione a più livelli in risposta alle richieste dei consumatori per l'argomento e il broker specificati. Questo parametro include tutte le partizioni dell'argomento che contribuiscono al traffico di trasferimento dati a valle sul broker specificato. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.
RemoteCopyBytesPerSec (RemoteBytesOutPerSec in v2.8.2.tiered)	Dopo aver creato un argomento, l'argomento è in fase di produzione/utilizzo.	Il numero di byte trasferiti nell'archiviazione a più livelli per l'argomento e il broker specificati. Questo parametro include tutte le partizioni dell'argomento che contribuiscono al traffico di trasferimento dati a monte sul broker specificato. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.
RemoteFetchErrorsPerSec (RemoteReadErrorPerSec in v2.8.2.tiered)	Dopo aver creato un argomento, l'argomento è in fase di	La percentuale di errori in risposta alle richieste di lettura che il broker specificato invia all'archiviazione a più livelli per recuperare i dati in risposta alle richieste dei consumatori in relazione all'argomento specificato. Questo parametro include tutte le partizioni dell'argo

Nome	Quando visibile	Descrizione
	produzione/ utilizzo.	mento che contribuiscono al traffico di trasferim ento dati a valle sul broker specificato. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.
`RemoteFetchRequest sPerSec (RemoteRe adRequestsPerSec in v2.8.2.tiered)`	Dopo aver creato un argomento, l'argomento è in fase di produzione/ utilizzo.	Il numero di richieste di lettura che il broker specificato invia all'archiviazione a più livelli per recuperare i dati in risposta alle richieste dei consumatori in relazione all'argomento specifica to. Questo parametro include tutte le partizioni dell'argomento che contribuiscono al traffico di trasferimento dati a valle sul broker specificato. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.
`RemoteCopyErrorsPe rSec (RemoteWr iteErrorPerSec in v2.8.2.tiered)`	Dopo aver creato un argomento, l'argomento è in fase di produzione/ utilizzo.	La percentuale di errori in risposta alle richieste di scrittura che il broker specificato invia all'archi viazione a più livelli per trasferire i dati a monte. Questo parametro include tutte le partizioni dell'argomento che contribuiscono al traffico di trasferimento dati a monte sul broker specifica to. Categoria: traffico e tassi di errore. Questo è un parametro KIP-405.

Monitoraggio del livello **`PER_TOPIC_PER_PARTITION`**

Quando imposti il livello di monitoraggio su `PER_TOPIC_PER_PARTITION`, ottieni i parametri descritti nella tabella seguente, oltre a tutti i parametri dei livelli `PER_TOPIC_PER_BROKER`, `PER_BROKER` e DEFAULT. Solo i parametri del livello `DEFAULT` sono gratuiti. I parametri in questa tabella hanno le seguenti dimensioni: gruppo di consumatori, argomento, partizione.

Parametri aggiuntivi disponibili a partire dal livello di monitoraggio **`PER_TOPIC_PER_PARTITION`**

Nome	Quando visibile	Descrizione
`EstimatedTimeLag`	Dopo che un gruppo di consumatori ha utilizzato un argomento.	Tempo stimato (in secondi) per eliminare il ritardo di offset della partizione.
`OffsetLag`	Dopo che un gruppo di consumatori ha utilizzato un argomento.	Ritardo del consumatore a livello di partizione nel numero di offset.

Visualizzazione dei parametri di Amazon MSK utilizzando CloudWatch

Puoi monitorare i parametri per Amazon MSK utilizzando la CloudWatch console, la riga di comando o l' CloudWatch API. Le procedure seguenti mostrano come accedere ai parametri utilizzando questi diversi metodi.

Per accedere alle metriche utilizzando la console CloudWatch

Accedi AWS Management Console e apri la CloudWatch console all'indirizzo https://console.aws.amazon.com/cloudwatch/.

1. Nel riquadro di navigazione, seleziona Parametri.
2. Scegli la scheda Tutti i parametri, quindi scegli AWS/Kafka.
3. Per visualizzare i parametri a livello di argomento, scegliere Topic, Broker ID, Cluster Name (Argomento, ID broker, Nome cluster); per parametri a livello di broker, scegliere Broker ID, Cluster Name (ID broker, Nome cluster); e per parametri a livello di cluster, scegliere Cluster Name (Nome cluster).
4. (Facoltativo) Nel riquadro grafico, seleziona una statistica e un periodo di tempo, quindi crea un CloudWatch allarme utilizzando queste impostazioni.

Per accedere alle metriche utilizzando il AWS CLI

Utilizza i comandi list-metrics e get-metric-statistics.

Per accedere alle metriche utilizzando la CLI CloudWatch

Utilizza i comandi mon-list-metrics e mon-get-stats.

Per accedere alle metriche utilizzando l'API CloudWatch

Utilizza le operazioni ListMetricsand GetMetricStatistics.

Monitoraggio del ritardo dei consumatori

Il monitoraggio del ritardo dei consumatori consente di identificare i consumatori lenti o bloccati che non tengono il passo con i dati più recenti disponibili su un argomento. Se necessario, puoi quindi intraprendere operazioni correttive, come il dimensionamento o il riavvio di tali consumatori. Per monitorare il ritardo dei consumatori, puoi utilizzare Amazon CloudWatch o open monitoring with Prometheus.

I parametri relativi al ritardo dei consumatori quantificano la differenza tra i dati più recenti scritti sui tuoi argomenti e i dati letti dalle tue applicazioni. Amazon MSK fornisce le seguenti metriche relative al ritardo dei consumatori, che puoi ottenere tramite Amazon CloudWatch o tramite il monitoraggio aperto con Prometheus:,,, e. `EstimatedMaxTimeLag` `EstimatedTimeLag` `MaxOffsetLag` `OffsetLag` `SumOffsetLag` Per ulteriori informazioni su questi parametri, consulta the section called “Metriche di Amazon MSK per il monitoraggio con CloudWatch”.

> **Note**
>
> Le metriche relative al ritardo dei consumatori sono visibili solo per i gruppi di consumatori in uno stato STABILE. Un gruppo di consumatori è STABILE dopo il completamento con successo del riequilibrio, garantendo che le partizioni siano distribuite uniformemente tra i consumatori.

Amazon MSK supporta i parametri di ritardo dei consumatori per i cluster con Apache Kafka 2.2.1 o una versione successiva.

Monitoraggio aperto con Prometheus

È possibile monitorare il cluster MSK con Prometheus, un sistema di monitoraggio open source per i dati dei parametri delle serie temporali. Puoi pubblicare questi dati su Servizio gestito da Amazon per Prometheus utilizzando la funzione di scrittura remota di Prometheus. Puoi anche utilizzare strumenti compatibili con parametri in formato Prometeo o strumenti che si integrano con Amazon MSK Open Monitoring, come Datadog, Lenses, New Relic e Sumo Logic. Il monitoraggio aperto è disponibile gratuitamente, ma per il trasferimento dei dati tra le zone di disponibilità vengono addebitati dei costi. Per informazioni su Prometheus, consulta la documentazione di Prometheus.

Creazione di un cluster Amazon MSK con monitoraggio aperto abilitato

Usando il AWS Management Console

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Nella sezione Monitoring (Monitoraggio), selezionare la casella di controllo accanto a Enable open monitoring with Prometheus (Abilita monitoraggio aperto con Prometheus).
3. Fornire le informazioni richieste in tutte le sezioni della pagina e rivedere tutte le opzioni disponibili.
4. Scegli Create cluster (Crea cluster).

Usando il AWS CLI

- Richiamare il comando create-cluster e specificarne l'opzione `open-monitoring`. Abilitare `JmxExporter`, `NodeExporter` o entrambi. Se si specifica `open-monitoring`, non è possibile disabilitare i due esportatori contemporaneamente.

Utilizzo dell'API

- Richiama l'CreateClusteroperazione e specifica`OpenMonitoring`. Abilitare `jmxExporter`, `nodeExporter` o entrambi. Se si specifica `OpenMonitoring`, non è possibile disabilitare i due esportatori contemporaneamente.

Abilitazione del monitoraggio aperto per un cluster Amazon MSK esistente

Per abilitare il monitoraggio aperto, assicurati che lo stato del cluster sia `ACTIVE`.

Usando il AWS Management Console

1. Accedi a e apri AWS Management Console la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/home?region=us-east-1#/home/.
2. Scegliere il nome del cluster da aggiornare. In questo modo si accede alla pagina dei dettagli del cluster.
3. Nella scheda Proprietà, scorri verso il basso per trovare la sezione Monitoraggio.
4. Scegli Modifica.
5. Selezionare la casella di controllo accanto a Enable open monitoring with Prometheus (Abilita monitoraggio aperto con Prometheus).
6. Seleziona Salvataggio delle modifiche.

Usando il AWS CLI

- Richiama il comando update-monitoring e specifica l'opzione `open-monitoring`. Abilitare `JmxExporter`, `NodeExporter` o entrambi. Se si specifica `open-monitoring`, non è possibile disabilitare i due esportatori contemporaneamente.

Utilizzo dell'API

- Richiama l'UpdateMonitoringoperazione e specifica`OpenMonitoring`. Abilitare `jmxExporter`, `nodeExporter` o entrambi. Se si specifica `OpenMonitoring`, non è possibile disabilitare i due esportatori contemporaneamente.

Impostazione di un host Prometheus su un'istanza Amazon EC2

1. Scaricare il server Prometheus da https://prometheus.io/download/#prometheus nell'istanza Amazon EC2.
2. Estrarre il file scaricato in una directory e passare a tale directory.
3. Creare un file denominato `prometheus.yml` con i seguenti contenuti:

```
# file: prometheus.yml
```

```
# my global config
global:
  scrape_interval:     60s

# A scrape configuration containing exactly one endpoint to scrape:
# Here it's Prometheus itself.
scrape_configs:
  # The job name is added as a label `job=<job_name>` to any timeseries scraped
 from this config.
  - job_name: 'prometheus'
    static_configs:
    # 9090 is the prometheus server port
    - targets: ['localhost:9090']
  - job_name: 'broker'
    file_sd_configs:
    - files:
      - 'targets.json'
```

4. Usa l'ListNodesoperazione per ottenere un elenco dei broker del tuo cluster.
5. Creare un file denominato `targets.json` con il seguente JSON. Sostituire *broker_dns_1*, *broker_dns_2* e il resto dei nomi DNS del broker con i nomi DNS ottenuti per i broker nella fase precedente. Includi tutti i broker ottenuti nel passaggio precedente. Amazon MSK utilizza la porta 11001 per JMX Exporter e la porta 11002 per Node Exporter.

   ZooKeeper mode targets.json

```
[
  {
    "labels": {
      "job": "jmx"
    },
    "targets": [
      "broker_dns_1:11001",
      "broker_dns_2:11001",
      .
      .
      .
      "broker_dns_N:11001"
    ]
  },
  {
    "labels": {
      "job": "node"
```

```
    },
    "targets": [
      "broker_dns_1:11002",
      "broker_dns_2:11002",
      .
      .
      .
      "broker_dns_N:11002"
    ]
  }
]
```

KRaft mode targets.json

```
[
  {
    "labels": {
      "job": "jmx"
    },
    "targets": [
      "broker_dns_1:11001",
      "broker_dns_2:11001",
      .
      .
      .
      "broker_dns_N:11001",
      "controller_dns_1:11001",
      "controller_dns_2:11001",
      "controller_dns_3:11001"
    ]
  },
  {
    "labels": {
      "job": "node"
    },
    "targets": [
      "broker_dns_1:11002",
      "broker_dns_2:11002",
      .
      .
      .
      "broker_dns_N:11002"
    ]
```

```
  }
]
```

> **Note**
>
> Per estrarre le metriche JMX dai controller Kraft, aggiungi i nomi DNS dei controller come destinazioni nel file JSON. Ad esempio: sostituendo `controller_dns_1` il nome DNS del controller con `controller_dns_1:11001` il nome DNS effettivo.

6. Per avviare il server Prometheus sull'istanza Amazon EC2, esegui il seguente comando nella directory in cui sono stati estratti i file Prometheus e salvati `prometheus.yml` e `targets.json`.

```
./prometheus
```

7. Individua l'indirizzo IP pubblico IPv4 dell'istanza Amazon EC2 in cui è stato eseguito Prometheus nel passaggio precedente. Questo indirizzo IP pubblico è necessario nella fase seguente.
8. Per accedere all'interfaccia utente Web di Prometheus, apri un browser che dispone dell'accesso all'istanza Amazon EC2 e passa a *`Prometheus-Instance-Public-IP`*`:9090`, dove *`Prometheus-Instance-Public-IP`* è l'indirizzo IP pubblico ottenuto nel passaggio precedente.

Parametri Prometheus

Tutti i parametri inviati da Apache Kafka a JMX sono accessibili tramite il monitoraggio aperto con Prometheus. Per informazioni sui parametri Apache Kafka, consulta la sezione relativa al monitoraggio nella documentazione di Apache Kafka. Oltre ai parametri di Apache Kafka, i parametri relativi al ritardo dei consumatori sono disponibili anche sulla porta 11001 con il nome JMX MBean `kafka.consumer.group:type=ConsumerLagMetrics`. Puoi anche utilizzare Prometheus Node Exporter per ottenere i parametri della CPU e del disco per i tuoi broker sulla porta 11002.

Archiviazione dei parametri Prometheus in Amazon Managed Service for Prometheus

Servizio gestito da Amazon per Prometheus è un servizio di monitoraggio e avviso compatibile con Prometheus che puoi utilizzare per monitorare i cluster Amazon MSK. È un servizio completamente

gestito che dimensiona automaticamente l'importazione, l'archiviazione, le query e gli avvisi dei parametri. Si integra inoltre con i servizi AWS di sicurezza per offrirti un accesso rapido e sicuro ai tuoi dati. È possibile utilizzare il linguaggio di query open source ProMQL per fare una query e creare avvisi relativi ai parametri.

Per ulteriori informazioni, consultare Guida introduttiva ad Amazon Managed Service for Prometheus.

Avvisi sulla capacità di archiviazione di Amazon MSK

Nei cluster con provisioning di Amazon MSK, scegli la capacità di archiviazione principale del cluster. L'esaurimento della capacità di archiviazione di un broker nel cluster con provisioning può influire sulla sua capacità di produrre e consumare dati, causando costosi tempi di inattività. Amazon MSK offre CloudWatch parametri per aiutarti a monitorare la capacità di storage del cluster. Inoltre, Amazon MSK invia automaticamente avvisi dinamici sulla capacità di archiviazione del cluster in modo da semplificare il rilevamento e la risoluzione dei problemi correlati. Gli avvisi sulla capacità di archiviazione includono raccomandazioni sulle misure a breve e a lungo termine necessarie per gestire la capacità di archiviazione del cluster. Dalla console Amazon MSK, puoi utilizzare i collegamenti rapidi all'interno degli avvisi per intraprendere immediatamente le operazioni consigliate.

Esistono due tipi di avvisi MSK sulla capacità di archiviazione: proattivi e correttivi.

- Gli avvisi proattivi sulla capacità di archiviazione ("Operazione richiesta") segnalano i potenziali problemi di archiviazione del cluster. Quando un broker in un cluster MSK ha utilizzato oltre il 60% o l'80% della sua capacità di archiviazione su disco, riceverai avvisi proattivi per il broker interessato.
- Gli avvisi correttivi relativi alla capacità di archiviazione ("Operazione critica richiesta") prevedono l'adozione di misure correttive per risolvere un problema critico del cluster quando uno dei broker del cluster MSK ha esaurito la capacità di archiviazione su disco.

Amazon MSK invia automaticamente questi avvisi alla console Amazon MSK,AWS Health Dashboard EventBridge, Amazon e ai contatti e-mail del tuo account. AWS Puoi anche configurare Amazon EventBridge per inviare questi avvisi a Slack o a strumenti come New Relic e Datadog.

Gli avvisi sulla capacità di archiviazione sono abilitati per impostazione predefinita per tutti i cluster MSK con provisioning e non possono essere disattivati. Questa funzionalità è supportata in tutte le regioni in cui è disponibile MSK.

Monitoraggio degli avvisi sulla capacità di archiviazione di Amazon MSK

Puoi verificare la presenza di avvisi sulla capacità di archiviazione in diversi modi:

- Accedi alla console Amazon MSK. Gli avvisi sulla capacità di archiviazione vengono visualizzati nel riquadro degli avvisi del cluster per 90 giorni. Gli avvisi contengono raccomandazioni e operazioni da eseguire con un solo clic per risolvere i problemi di capacità di archiviazione su disco.
- Usa ListClusterse le API ListClustersV2 o DescribeClusterV2 per visualizzare tutti `CustomerActionStatus` gli avvisi relativi a un cluster. DescribeCluster
- Vai alla AWS Health Dashboard per visualizzare gli avvisi di MSK e di altri AWS servizi.
- Configura AWS Health API e Amazon EventBridge per indirizzare le notifiche di avviso a piattaforme di terze parti come Datadog e NewRelic Slack.

Utilizzo LinkedIn del Cruise Control per Apache Kafka con Amazon MSK

Puoi utilizzare LinkedIn Cruise Control per ribilanciare il cluster Amazon MSK, rilevare e correggere anomalie e monitorare lo stato e l'integrità del cluster.

Download e compilazione di Cruise Control

1. Crea un'istanza Amazon EC2 nello stesso Amazon VPC del cluster Amazon MSK.
2. Installa Prometheus sull'istanza Amazon EC2 che hai creato nel passaggio precedente. Prendi nota dell'IP privato e della porta. Il numero di porta predefinito è 9090. Per informazioni su come configurare Prometheus per aggregare i parametri per un cluster, consulta la pagina the section called “Monitoraggio aperto con Prometheus”.
3. Scarica Cruise Control sull'istanza Amazon EC2. In alternativa, se preferisci, puoi utilizzare un'istanza Amazon EC2 separata per Cruise Control. Per un cluster con Apache Kafka versione 2.4.*, usa la versione 2.4.* di Cruise Control più recente. Se il tuo cluster ha una versione di Apache Kafka precedente alla 2.4.*, utilizza la versione 2.0.* di Cruise Control più recente.
4. Decomprimi il file Cruise Control, quindi vai alla cartella decompressa.
5. Esegui il comando seguente per installare git.

   ```
   sudo yum -y install git
   ```

6. Esegui il comando seguente per inizializzare il repository locale. Sostituisci *Your-Cruise-Control-Folder con il nome della cartella* corrente (la cartella che hai ottenuto quando hai decompresso il download di cruise control).

   ```
   git init && git add . && git commit -m "Init local repo." && git tag -a Your-Cruise-Control-Folder -m "Init local version."
   ```

7. Esegui il comando seguente per creare il codice sorgente.

   ```
   ./gradlew jar copyDependantLibs
   ```

Configurazione ed esecuzione di Cruise Control

1. Apporta le seguenti modifiche al file `config/cruisecontrol.properties`. Sostituisci la stringa bootstrap servers e bootstrap-brokers di esempio con i valori del tuo cluster. Per recuperare queste stringhe per il cluster, puoi consultare i dettagli del cluster nella console. In alternativa, puoi utilizzare le operazioni GetBootstrapBrokerse DescribeClusterAPI o i loro equivalenti CLI.

```
# If using TLS encryption, use 9094; use 9092 if using plaintext
bootstrap.servers=b-1.test-cluster.2skv42.c1.kafka.us-
east-1.amazonaws.com:9094,b-2.test-cluster.2skv42.c1.kafka.us-
east-1.amazonaws.com:9094,b-3.test-cluster.2skv42.c1.kafka.us-
east-1.amazonaws.com:9094

# SSL properties, needed if cluster is using TLS encryption
security.protocol=SSL
ssl.truststore.location=/home/ec2-user/kafka.client.truststore.jks

# Use the Prometheus Metric Sampler
metric.sampler.class=com.linkedin.kafka.cruisecontrol.monitor.sampling.prometheus.Prometheu

# Prometheus Metric Sampler specific configuration
prometheus.server.endpoint=1.2.3.4:9090 # Replace with your Prometheus IP and port

# Change the capacity config file and specify its path; details below
capacity.config.file=config/capacityCores.json
```

2. Modifica il file `config/capacityCores.json` per specificare le dimensioni corrette del disco, i core della CPU e i limiti di ingresso/uscita della rete. È possibile utilizzare l'operazione DescribeClusterAPI (o il suo equivalente CLI) per ottenere la dimensione del disco. Per i core della CPU e i limiti di ingresso/uscita di rete, consulta la pagina Tipi di istanza Amazon EC2.

```
{
  "brokerCapacities": [
    {
      "brokerId": "-1",
      "capacity": {
        "DISK": "10000",
        "CPU": {
          "num.cores": "2"
        },
        "NW_IN": "5000000",
```

```
        "NW_OUT": "5000000"
      },
      "doc": "This is the default capacity. Capacity unit used for disk is in MB,
 cpu is in number of cores, network throughput is in KB."
    }
  ]
}
```

3. Facoltativamente, puoi installare l'interfaccia utente di Cruise Control. Per scaricarla, consulta la pagina Setting Up Cruise Control Frontend.
4. Esegui il comando seguente per avviar Cruise Control. Prendi in considerazione l'utilizzo di uno strumento come `screen` o `tmux` per mantenere aperta una sessione di lunga durata.

```
<path-to-your-kafka-installation>/bin/kafka-cruise-control-start.sh config/
cruisecontrol.properties 9091
```

5. Utilizza le API o l'interfaccia utente di Cruise Control per assicurarti che Cruise Control disponga dei dati di carico del cluster e che fornisca suggerimenti per il ribilanciamento. Potrebbero trascorrere alcuni minuti prima di ottenere una finestra di parametri valida.

Modello di distribuzione automatizzata di Cruise Control per Amazon MSK

Puoi anche utilizzare questo CloudFormation modello per implementare facilmente Cruise Control e Prometheus per ottenere informazioni più approfondite sulle prestazioni del tuo cluster Amazon MSK e ottimizzare l'utilizzo delle risorse.

Caratteristiche principali:

- Provisioning automatico di un'istanza Amazon EC2 con Cruise Control e Prometheus preconfigurati.
- Support per il cluster con provisioning di Amazon MSK.
- Autenticazione flessibile con PlainText e IAM.
- Nessuna dipendenza da Zookeeper per Cruise Control.
- Personalizza facilmente gli obiettivi Prometheus, le impostazioni della capacità del Cruise Control e altre configurazioni fornendo i tuoi file di configurazione memorizzati in un bucket Amazon S3.

Quota di Amazon MSK

Il tuo AWS account ha quote predefinite per Amazon MSK. Salvo diversa indicazione, ogni quota per account è specifica per regione all'interno dell'account. AWS

Quota di Amazon MSK

- Fino a 90 broker per account. 30 broker per cluster modale. 60 broker per cluster ZooKeeper in modalità Kraft. Per richiedere una quota più elevata, vai al Support Center della AWS console e crea un caso di supporto.
- Un minimo di 1 GiB di archiviazione per broker.
- Un massimo di 16.384 GiB di archiviazione per broker.
- Un cluster che utilizza the section called “Controllo degli accessi IAM” può avere fino a 3.000 connessioni TCP per broker in un dato momento. Per aumentare questo limite, puoi modificare la `listener.name.client_iam.max.connections` o la proprietà di `listener.name.client_iam_public.max.connections` configurazione utilizzando l' AlterConfig API Kafka o lo `kafka-configs.sh` strumento. È importante notare che impostare una delle due proprietà su un valore elevato può comportare l'indisponibilità.
- Limiti sulle connessioni TCP. Con l'opzione Connection Rate Burst abilitata, MSK consente 100 connessioni al secondo. L'eccezione è il tipo di istanza kafka.t3.small, a cui sono consentite 4 connessioni al secondo con i burst di velocità di connessione abilitati. I cluster più vecchi che non hanno abilitato i burst di velocità di connessione avranno la funzionalità abilitata automaticamente quando il cluster verrà aggiornato.

  Per gestire i tentativi di connessione non riusciti, puoi impostare il parametro di configurazione `reconnect.backoff.ms` sul lato client. Ad esempio, se desideri che un client ritenti la connessione dopo 1 secondo, imposta `reconnect.backoff.ms` su 1.000. Per ulteriori informazioni, consulta la sezione reconnect.backoff.ms nella documentazione di Apache Kafka.
- Fino a 100 configurazioni per account. Per richiedere una modifica della quota, visita il Centro assistenza della console AWS e crea un ticket.
- Un massimo di 50 revisioni per configurazione.
- Per aggiornare la configurazione o la versione Apache Kafka di un cluster MSK, assicurati innanzitutto che il numero di partizioni per broker sia inferiore ai limiti descritti in the section called “ Dimensionamento corretto del cluster: numero di partizioni per broker”.

Quote del replicatore MSK

- Un massimo di 15 replicatori MSK per account.
- MSK Replicator replica solo fino a 750 argomenti in ordine ordinato. Se è necessario replicare più argomenti, si consiglia di creare un Replicator separato. Vai al Support Center della AWS console e crea un caso di supporto se hai bisogno di supporto per più di 750 argomenti per Replicator. È possibile monitorare il numero di argomenti replicati utilizzando la metrica TopicCount "».
- Una velocità di trasmissione effettiva di ingresso massima di 1 GB al secondo per replicatore MSK. Per richiedere una quota più elevata, vai al Support Center della AWS console e crea un caso di supporto.
- Dimensione record MSK Replicator: dimensione massima del record di 10 MB (message.max.bytes). Per richiedere una quota più elevata, vai al Support Center della AWS console e crea un caso di supporto.

Quota di MSK Serverless

> **Note**
>
> In caso di problemi con i limiti di quota, contatta l' AWS assistenza creando una richiesta di supporto.

I limiti si intendono per cluster, salvo diversa indicazione.

Dimensione	Quota	Risultato della violazione della quota
Velocità di trasmissione effettiva massima in ingresso	200 Mb/s	Rallentamento con limitazione della larghezza di banda della rete prolungata in risposta
Velocità di trasmissione effettiva massima in uscita	400 Mb/s	Rallentamento con limitazione della larghezza di banda della rete prolungata in risposta

Dimensione	Quota	Risultato della violazione della quota
Durata massima di conservazione	Illimitato	N/D
Numero massimo di connessioni client	3000	Chiusura della connessione
Numero massimo di tentativi di connessione	100 al secondo	Chiusura della connessione
Dimensione massima del messaggio	8 MB	La richiesta ha esito negativo con ErrorCode: INVALID_REQUEST
Velocità massima di richieste	15.000 al secondo	Rallentamento con limitazione della larghezza di banda della rete prolungata in risposta
Velocità massima di richieste delle API di gestione degli argomenti	2 al secondo	Rallentamento con limitazione della larghezza di banda della rete prolungata in risposta
Numero massimo di byte di recupero per richiesta	55 MB	La richiesta fallisce con: INVALID_REQUEST ErrorCode
Numero massimo di gruppi di consumatori	500	JoinGroup la richiesta fallisce
Numero massimo di partizioni (leader)	2.400 per argomenti non compattati. 120 per argomenti compattati. Per richiedere un aggiustamento della quota, vai al Support Center della AWS console e crea una richiesta di supporto.	La richiesta fallisce con ErrorCode: INVALID_REQUEST

Dimensione	Quota	Risultato della violazione della quota
Velocità massima di creazione ed eliminazione delle partizioni	250 in 5 minuti	La richiesta ha esito negativo con: THROUGHPUT_QUOTA_EXCEEDED ErrorCode
Velocità di trasmissione effettiva massima in ingresso per partizione	5 Mb/s	Rallentamento con limitazione della larghezza di banda della rete prolungata in risposta
Velocità di trasmissione effettiva massima in uscita per partizione	10 Mb/s	Rallentamento con limitazione della larghezza di banda della rete prolungata in risposta
Dimensione massima della partizione (per argomenti compatti)	250 GB	La richiesta non riesce con ErrorCode: THROUGHPUT_QUOTA_EXCEEDED
Numero massimo di VPC client per cluster serverless	5	
Numero massimo di cluster serverless per account	10. Per richiedere un aggiustamento della quota, vai al Support Center della AWS console e crea una richiesta di supporto.	

Quota di MSK Connect

- Fino a 100 plug-in personalizzati.
- Fino a 100 configurazioni di worker.
- Fino a 60 worker di connessione. Se un connettore è configurato in modo da avere una capacità con dimensionamento automatico, MSK Connect utilizza il numero massimo di worker che il connettore è configurato per avere al fine di calcolare la quota per l'account.
- Fino a 10 worker per connettore.

Per richiedere una quota più elevata per MSK Connect, vai al Support Center della AWS console e crea un caso di supporto.

Risorse Amazon MSK

Il termine risorse ha due significati in Amazon MSK, a seconda del contesto. Nel contesto delle API, una risorsa è una struttura sulla quale è possibile richiamare un'operazione. Per un elenco di queste risorse e delle operazioni che è possibile richiamare su di esse, consulta la sezione Resources nella documentazione di riferimento dell'API di Amazon MSK. Nel contesto di the section called “Controllo degli accessi IAM”, una risorsa è un'entità a cui è possibile consentire o rifiutare l'accesso, come definito nella sezione the section called “Risorse”.

Integrazioni di MSK

Questa sezione fornisce riferimenti alle AWS funzionalità che si integrano con Amazon MSK.

Argomenti

Connettore Amazon Athena per Amazon MSK

Il connettore Amazon Athena per Amazon MSK consente ad Amazon Athena di eseguire query SQL sugli argomenti di Apache Kafka. Utilizza questo connettore per visualizzare gli argomenti di Apache Kafka come tabelle e i messaggi come righe in Athena.

Per ulteriori informazioni, consulta la pagina Amazon Athena MSK Connector nella Guida per l'utente di Amazon Athena.

Importazione dei dati in streaming con Amazon Redshift

Amazon Redshift supporta l'importazione dei dati in streaming da Amazon MSK. La funzionalità di importazione dei dati in streaming di Amazon Redshift fornisce l'importazione a bassa latenza e ad alta velocità dei dati in streaming da Amazon MSK in una vista materializzata di Amazon Redshift. Poiché non richiede la gestione temporanea dei dati in Amazon S3, Amazon Redshift può importare dati in streaming a una latenza inferiore e a un costo di archiviazione ridotto. Puoi configurare l'importazione dei dati in streaming di Amazon Redshift su un cluster Amazon Redshift utilizzando le istruzioni SQL per autenticarti e connetterti a un argomento Amazon MSK.

Per ulteriori informazioni, consulta la pagina Streaming ingestion nella Guida per gli sviluppatori di database di Amazon Redshift.

Firehose

Amazon MSK si integra con Firehose per fornire una soluzione serverless e senza codice per distribuire flussi dai cluster Apache Kafka ai data lake Amazon S3. Firehose è un servizio di

streaming di estrazione, trasformazione e caricamento (ETL) che legge i dati dagli argomenti di Amazon MSK Kafka, esegue trasformazioni come la conversione in Parquet e aggrega e scrive i dati su Amazon S3. Con pochi clic dalla console, puoi configurare uno stream Firehose da leggere da un argomento di Kafka e inviarlo a una posizione S3. Non richiede la scrittura di codice, applicazioni di connessione né provisioning di risorse. Firehose si ridimensiona automaticamente in base alla quantità di dati pubblicati sull'argomento Kafka e paghi solo per i byte acquisiti da Kafka.

Per ulteriori informazioni su questa funzionalità, consulta le seguenti risorse.

- Scrittura su Kinesis Data Firehose utilizzando Amazon MSK - Amazon Kinesis Data Firehose nella Amazon Data Firehose Developer Guide
- Post del blog: Amazon MSK Introduces Managed Data Delivery from Apache Kafka to Your Data Lake
- Laboratorio: consegna ad Amazon S3 tramite Firehose

Accesso ad Amazon EventBridge Pipes tramite la console Amazon MSK

Amazon EventBridge Pipes collega le sorgenti alle destinazioni. Le pipe sono destinate point-to-point alle integrazioni tra sorgenti e destinazioni supportate, con supporto per trasformazioni e arricchimenti avanzati. EventBridge Le pipe offrono un modo altamente scalabile per connettere il tuo cluster Amazon MSK a AWS servizi come Step Functions, Amazon SQS e API Gateway, nonché ad applicazioni SaaS (Software as a Service) di terze parti come Salesforce.

Per configurare una pipe, scegli l'origine, aggiungi filtri opzionali, definisci l'arricchimento opzionale e scegli la destinazione per i dati dell'evento.

Nella pagina dei dettagli di un cluster Amazon MSK, puoi visualizzare le pipe che utilizzano quel cluster come origine. Da lì, puoi anche:

- Avvia la console per visualizzare i dettagli delle pipe. EventBridge
- Avvia la EventBridge console per creare una nuova pipe con il cluster come origine.

Per ulteriori informazioni sulla configurazione di un cluster Amazon MSK come sorgente pipe, consulta Amazon Managed Streaming for Apache Kafka cluster come sorgente nella Amazon User Guide. EventBridge Per ulteriori informazioni su Pipes in generale, consulta EventBridge Pipes. EventBridge

Per accedere alle EventBridge pipe per un determinato cluster Amazon MSK

1. Apri la console Amazon MSK e seleziona Cluster.
2. Seleziona un cluster.
3. Nella pagina Dettagli del cluster, scegli la scheda Integrazione.

   La scheda Integrazione include un elenco di tutte le pipe attualmente configurate per utilizzare il cluster selezionato come origine, tra cui:

   - nome della pipe
   - stato corrente
   - destinazione della pipe
   - quando la pipe è stata modificata l'ultima volta
4. Gestisci le pipe per il tuo cluster Amazon MSK come desideri:

   Accesso a dettagli aggiuntivi su una pipe

   - Scegli la pipe.

     Verrà avviata la pagina dei dettagli di Pipe della EventBridge console.

   Creazione di una nuova pipe

   - Scegli Connetti il cluster Amazon MSK alla pipe.

     Verrà avviata la pagina Create pipe della EventBridge console, con il cluster Amazon MSK specificato come origine pipe. Per ulteriori informazioni, consulta Creating an EventBridge pipe nella Amazon EventBridge User Guide.

   - Puoi creare una pipe per un cluster anche dalla pagina Cluster. Seleziona il cluster e, dal menu Azioni, seleziona Create EventBridge Pipe.

Versioni di Apache Kafka

Quando si crea un cluster Amazon MSK, specifica quale versione di Apache Kafka desideri utilizzare. Puoi inoltre aggiornare la versione di Apache Kafka di un cluster esistente. Gli argomenti del capitolo aiutano a comprendere le tempistiche per il supporto delle versioni di Kafka e i suggerimenti per le migliori pratiche.

Argomenti

Versioni di Apache Kafka supportate

Streaming gestito da Amazon per Apache Kafka (Amazon MSK) supporta le seguenti versioni di Apache Kafka e Amazon MSK. La community di Apache Kafka fornisce circa 12 mesi di supporto per una versione successiva alla data di rilascio. Per maggiori dettagli, consulta la politica EOL (end of life) di Apache Kafka.

Versioni di Apache Kafka supportate

Versione Apache Kafka	Data di rilascio di MSK	Data di fine del supporto
1.1.1	--	2024-06-05
2.1.0	--	2024-06-05
2.2.1	31-07-2019	2024-06-08
2.3.1	19-12-2019	2024-06-08
24.1	2020-04-02	2024-06-08
2,41.1	2020-09-09	2024-06-08
2.5.1	2020-09-30	2024-06-08
2,6,0	2020-10-21	2024-09-11
2,6,1	2021-01-19	2024-09-11

Versione Apache Kafka	Data di rilascio di MSK	Data di fine del supporto
2,6,2	2021-04-29	2024-09-11
2,6,3	2021-12-21	2024-09-11
2,7,0	2020-12-29	2024-09-11
2,7,1	2021-05-25	2024-09-11
2,7,2	2021-12-21	2024-09-11
2,80	--	2024-09-11
28,1	2022-10-28	2024-09-11
2.8.2 livelli	2022-10-28	Da annunciare
3.1.1	2022-06-22	2024-09-11
32,0	2022-06-22	2024-09-11
3,31	2022-10-26	2024-09-11
3,32	-02	2024-09-11
3,40	2023-05-04	2025-06-17
3.5.1 (consigliato)	2023-09-26	--
3,6,0	2023-11-16	--
3,7. x	2024-05-29	--

Per ulteriori informazioni sulla politica di supporto delle versioni di Amazon MSK, consultaPolitica di supporto delle versioni di Amazon MSK.

Apache Kafka versione 3.7.x (con storage su più livelli pronto per la produzione)

La versione 3.7.x di Apache Kafka su MSK include il supporto per Apache Kafka versione 3.7.0. È possibile creare cluster o aggiornare i cluster esistenti per utilizzare la nuova versione 3.7.x. Con questa modifica nella denominazione delle versioni, non è più necessario adottare versioni di patch fix più recenti come la 3.7.1 quando vengono rilasciate dalla community di Apache Kafka. Amazon MSK aggiornerà automaticamente la versione 3.7.x per supportare le future versioni delle patch non appena saranno disponibili. In questo modo puoi sfruttare la sicurezza e le correzioni di bug disponibili tramite le versioni patch fix senza attivare un aggiornamento della versione. Queste versioni di patch fix rilasciate da Apache Kafka non compromettono la compatibilità delle versioni e puoi trarre vantaggio dalle nuove versioni di patch fix senza preoccuparti degli errori di lettura o scrittura delle applicazioni client. Assicurati che gli strumenti di automazione dell'infrastruttura, ad esempio CloudFormation, siano aggiornati per tenere conto di questa modifica nella denominazione delle versioni.

Amazon MSK ora supporta la modalità KRAFT (Apache Kafka Raft) nella versione 3.7.x di Apache Kafka. Su Amazon MSK, come per i ZooKeeper nodi, i controller Kraft sono inclusi senza costi aggiuntivi e non richiedono alcuna configurazione o gestione aggiuntiva da parte dell'utente. Ora puoi creare cluster in modalità KRAFT o in modalità Apache Kafka versione ZooKeeper 3.7.x. In modalità Kraft, puoi aggiungere fino a 60 broker per ospitare più partizioni per cluster, senza richiedere un aumento del limite, rispetto alla quota di 30 broker sui cluster basati su Zookeeper. Per saperne di più su KRAFT su MSK, consulta la modalità KRAFT.

La versione 3.7.x di Apache Kafka include anche diverse correzioni di bug e nuove funzionalità che migliorano le prestazioni. I miglioramenti principali includono le ottimizzazioni di Leader Discovery per i client e le opzioni di ottimizzazione del log Segment Flush. Per un elenco completo dei miglioramenti e delle correzioni di bug, consultate le note di rilascio di Apache Kafka per la versione 3.7.0.

Apache Kafka versione 3.6.0 (con archiviazione a più livelli pronta per la produzione)

Per informazioni su Apache Kafka versione 3.6.0 (con archiviazione a più livelli pronta per la produzione), consulta le relative note di rilascio sul sito dei download di Apache Kafka.

Per motivi di stabilità, Amazon MSK continuerà a utilizzare e gestire ZooKeeper per la gestione del quorum in questa versione.

Amazon MSK versione 3.5.1

Amazon Managed Streaming for Apache Kafka (Amazon MSK) ora supporta la versione 3.5.1 di Apache Kafka per cluster nuovi ed esistenti. Apache Kafka 3.5.1 include diverse correzioni di bug e nuove funzionalità che migliorano le prestazioni. Le caratteristiche principali includono l'introduzione di una nuova assegnazione delle partizioni compatibile con i rack per i consumatori. Amazon MSK continuerà a utilizzare e gestire Zookeeper per la gestione del quorum in questa versione. Per un elenco completo dei miglioramenti e delle correzioni di bug, consulta le note di rilascio di Apache Kafka per la versione 3.5.1.

Per informazioni su Apache Kafka versione 3.5.1, consulta le relative note di rilascio sul sito dei download di Apache Kafka.

Amazon MSK versione 3.4.0

Amazon Managed Streaming for Apache Kafka (Amazon MSK) ora supporta Apache Kafka versione 3.4.0 per cluster nuovi ed esistenti. Apache Kafka 3.4.0 include diverse correzioni di bug e nuove funzionalità che migliorano le prestazioni. Le funzionalità principali includono una correzione per migliorare la stabilità da recuperare dalla replica più vicina. Amazon MSK continuerà a utilizzare e gestire Zookeeper per la gestione del quorum in questa versione. Per un elenco completo dei miglioramenti e delle correzioni di bug, consulta le note di rilascio di Apache Kafka per la versione 3.4.0.

Per informazioni su Apache Kafka versione 3.4.0, consulta le relative note di rilascio sul sito dei download di Apache Kafka.

Amazon MSK versione 3.3.2

Amazon Managed Streaming for Apache Kafka (Amazon MSK) ora supporta la versione 3.3.2 di Apache Kafka per cluster nuovi ed esistenti. Apache Kafka 3.3.2 include diverse correzioni di bug e nuove funzionalità che migliorano le prestazioni. Le funzionalità principali includono una correzione per migliorare la stabilità da recuperare dalla replica più vicina. Amazon MSK continuerà a utilizzare e gestire Zookeeper per la gestione del quorum in questa versione. Per un elenco completo dei miglioramenti e delle correzioni di bug, consulta le note di rilascio di Apache Kafka per la versione 3.3.2.

Per informazioni su Apache Kafka versione 3.3.2, consulta le relative note di rilascio sul sito dei download di Apache Kafka.

Amazon MSK versione 3.3.1

Amazon Managed Streaming for Apache Kafka (Amazon MSK) ora supporta la versione 3.3.1 di Apache Kafka per cluster nuovi ed esistenti. Apache Kafka 3.3.1 include diverse correzioni di bug e nuove funzionalità che migliorano le prestazioni. Alcune delle funzionalità principali includono miglioramenti alle metriche e al partizionatore. Per motivi di stabilità, Amazon MSK continuerà a utilizzare e gestire ZooKeeper per la gestione del quorum in questa versione. Per un elenco completo dei miglioramenti e delle correzioni di bug, consultate le note di rilascio di Apache Kafka per la versione 3.3.1.

Per informazioni su Apache Kafka versione 3.3.1, consulta le relative note di rilascio sul sito dei download di Apache Kafka.

Amazon MSK versione 3.1.1

Amazon Managed Streaming for Apache Kafka (Amazon MSK) ora supporta le versioni 3.1.1 e 3.2.0 di Apache Kafka per cluster nuovi ed esistenti. Apache Kafka 3.1.1 e Apache Kafka 3.2.0 includono diverse correzioni di bug e nuove funzionalità che migliorano le prestazioni. Alcune delle funzionalità principali includono miglioramenti alle metriche e l'uso degli ID degli argomenti. MSK continuerà a utilizzare e gestire Zookeeper per la gestione del quorum in questa versione per motivi di stabilità. Per un elenco completo dei miglioramenti e delle correzioni di bug, consultate le note di rilascio di Apache Kafka per 3.1.1 e 3.2.0.

Per informazioni sulle versioni 3.1.1 e 3.2.0 di Apache Kafka, consultate le relative note di rilascio 3.2.0 e le note di rilascio 3.1.1 sul sito di download di Apache Kafka.

Archiviazione a più livelli Amazon MSK versione 2.8.2.tiered

Questa versione è una versione solo per Amazon MSK di Apache Kafka versione 2.8.2 ed è compatibile con i client open source Apache Kafka.

La versione 2.8.2.tiered contiene funzionalità di archiviazione a più livelli compatibili con le API introdotte in KIP-405 per Apache Kafka. Per ulteriori informazioni sulla funzionalità di archiviazione a più livelli di Amazon MSK, consulta la sezione Archiviazione a più livelli.

Apache Kafka versione 2.5.1

La versione 2.5.1 di Apache Kafka include diverse correzioni di bug e nuove funzionalità, tra cui la crittografia in transito per Apache e i client di amministrazione. ZooKeeper Amazon MSK fornisce ZooKeeper endpoint TLS, che possono essere interrogati durante l'operazione. DescribeCluster

L'output dell' DescribeClusteroperazione include il `ZookeeperConnectStringTls` nodo, che elenca gli endpoint TLS zookeeper.

L'esempio seguente mostra il nodo `ZookeeperConnectStringTls` della risposta per l'operazione `DescribeCluster`:

```
"ZookeeperConnectStringTls": "z-3.awskafkatutorialc.abcd123.c3.kafka.us-
east-1.amazonaws.com:2182,z-2.awskafkatutorialc.abcd123.c3.kafka.us-
east-1.amazonaws.com:2182,z-1.awskafkatutorialc.abcd123.c3.kafka.us-
east-1.amazonaws.com:2182"
```

Per informazioni sull'utilizzo della crittografia TLS con ZooKeeper, consulta la sezione Utilizzo della sicurezza TLS con Apache ZooKeeper.

Per ulteriori informazioni su Apache Kafka versione 2.5.1, consulta le relative note di rilascio sul sito dei download di Apache Kafka.

Versione di correzione dei bug Amazon MSK 2.4.1.1

Questa versione è una versione di correzione dei bug di Apache Kafka 2.4.1 disponibile solo per Amazon MSK. Questa versione di correzione contiene una correzione per KAFKA-9752, un problema raro che causa il continuo ribilanciamento dei gruppi di consumatori e la permanenza nello stato `PreparingRebalance`. Questo problema riguarda i cluster che eseguono Apache Kafka versioni 2.3.1 e 2.4.1. Questa versione contiene una correzione prodotta dalla comunità disponibile nella versione 2.5.0 di Apache Kafka.

> **Note**
>
> I cluster Amazon MSK che eseguono la versione 2.4.1.1 sono compatibili con qualsiasi client Apache Kafka compatibile con la versione 2.4.1 di Apache Kafka.

Se preferisci usare Apache Kafka 2.4.1, ti consigliamo di utilizzare la versione 2.4.1.1 di correzione dei bug MSK per i nuovi cluster Amazon MSK. Per incorporare questa correzione, puoi aggiornare i cluster esistenti che eseguono Apache Kafka versione 2.4.1 a questa versione. Per informazioni sull'aggiornamento di un cluster esistente, consulta la sezione Aggiornamento della versione di Apache Kafka.

Per risolvere questo problema senza aggiornare il cluster alla versione 2.4.1.1, consulta la sezione Gruppo di consumatori bloccato nello stato `PreparingRebalance` della guida Risoluzione dei problemi relativi al cluster Amazon MSK.

Apache Kafka versione 2.4.1 (usa invece 2.4.1.1)

> **Note**
>
> Non è più possibile creare un cluster MSK con la versione 2.4.1 di Apache Kafka. In alternativa, è possibile utilizzare Versione di correzione dei bug Amazon MSK 2.4.1.1 con client compatibili con la versione 2.4.1 di Apache Kafka. Se disponi già di un cluster MSK con Apache Kafka versione 2.4.1, ti consigliamo di aggiornarlo per utilizzare invece la versione 2.4.1.1 di Apache Kafka.

KIP-392 è una delle principali proposte di miglioramento di Kafka incluse nella versione 2.4.1 di Apache Kafka. Questo miglioramento consente ai consumatori di recuperare dati dalla replica più vicina. Per utilizzare questa caratteristica, imposta `client.rack` nelle proprietà consumatore sull'ID della zona di disponibilità del consumatore. Un esempio di ID di zona di disponibilità è `use1-az1`. Amazon MSK imposta `broker.rack` sugli ID delle zone di disponibilità dei broker. Inoltre, devi impostare la proprietà di configurazione `replica.selector.class` su `org.apache.kafka.common.replica.RackAwareReplicaSelector`, che è un'implementazione di consapevolezza rack fornita da Apache Kafka.

Quando utilizzi questa versione di Apache Kafka, i parametri nel livello di monitoraggio `PER_TOPIC_PER_BROKER` vengono visualizzati solo dopo che i valori diventano diversi da zero per la prima volta. Per ulteriori informazioni, consulta the section called “Monitoraggio del livello `PER_TOPIC_PER_BROKER`”.

Per informazioni su come trovare gli ID delle zone di disponibilità, consulta AZ IDs for Your Resource nella guida per l'utente. AWS Resource Access Manager

Per informazioni sull'impostazione delle proprietà di configurazione, consulta Configurazione.

Per ulteriori informazioni su KIP-392, consulta Allow Consumers to Fetch from Closest Replica nelle pagine di Confluence.

Per ulteriori informazioni su Apache Kafka versione 2.4.1, consulta le relative note di rilascio sul sito dei download di Apache Kafka.

Supporto per la versione di Amazon MSK

Questo argomento descrive Politica di supporto delle versioni di Amazon MSK e la procedura perAggiornamento della versione di Apache Kafka. Se stai aggiornando la tua versione di Kafka, segui le migliori pratiche descritte in. Le migliori pratiche per gli aggiornamenti delle versioni

Politica di supporto delle versioni di Amazon MSK

Questa sezione descrive la politica di supporto per le versioni di Kafka supportate da Amazon MSK.

- Tutte le versioni di Kafka sono supportate fino al raggiungimento della data di fine del supporto. Per informazioni dettagliate sulle date di fine del supporto, consulta. Versioni di Apache Kafka supportate Aggiorna il tuo cluster MSK alla versione di Kafka consigliata o alla versione successiva prima della data di fine del supporto. Per dettagli sull'aggiornamento della versione di Apache Kafka, consulta. Aggiornamento della versione di Apache Kafka Un cluster che utilizza una versione di Kafka dopo la data di fine del supporto viene aggiornato automaticamente alla versione Kafka consigliata.
- MSK eliminerà gradualmente il supporto per i cluster di nuova creazione che utilizzano versioni di Kafka con date di fine supporto pubblicate.

Aggiornamento della versione di Apache Kafka

Ora è possibile aggiornare un cluster MSK esistente a una versione più recente di Apache Kafka. Non puoi aggiornarlo a una versione precedente. Quando aggiorni la versione di Apache Kafka di un cluster MSK, controlla anche il software lato client per assicurarti che la versione consenta di utilizzare le funzionalità della nuova versione Apache Kafka del cluster. Amazon MSK aggiorna soltanto il software del server. Non aggiorna i clienti.

Per informazioni su come rendere un cluster altamente disponibile durante un aggiornamento, consulta the section called “Creazione di cluster a disponibilità elevata”.

> **Important**
>
> Non è possibile aggiornare la versione di Apache Kafka per un cluster MSK che supera i limiti descritti nella pagina the section called “ Dimensionamento corretto del cluster: numero di partizioni per broker”.

Aggiornamento della versione di Apache Kafka utilizzando il AWS Management Console

1. Apri la console Amazon MSK all'indirizzo https://console.aws.amazon.com/msk/.
2. Scegli il cluster MSK per il quale desideri aggiornare la versione di Apache Kafka.
3. Nella scheda Proprietà, scegli Aggiorna nella sezione relativa alla versione di Apache Kafka.

Aggiornamento della versione di Apache Kafka utilizzando il AWS CLI

1. Esegui il comando seguente, sostituendolo *ClusterArn* con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

   ```
   aws kafka get-compatible-kafka-versions --cluster-arn ClusterArn
   ```

   L'output di questo comando include un elenco delle versioni di Apache Kafka a cui è possibile aggiornare il cluster. Il risultato sembra l'esempio seguente.

   ```
   {
       "CompatibleKafkaVersions": [
           {
               "SourceVersion": "2.2.1",
               "TargetVersions": [
                   "2.3.1",
                   "2.4.1",
                   "2.4.1.1",
                   "2.5.1"
               ]
           }
       ]
   }
   ```

2. Esegui il comando seguente, sostituendolo *ClusterArn* con l'Amazon Resource Name (ARN) che hai ottenuto quando hai creato il cluster. Se non disponi dell'ARN per il cluster, puoi trovarlo elencando tutti i cluster. Per ulteriori informazioni, consulta the section called “Elencazione dei cluster”.

Sostituisci *Current-Cluster-Version* con la versione corrente del cluster. Perché *TargetVersion*puoi specificare una qualsiasi delle versioni di destinazione dall'output del comando precedente.

> **Important**
>
> Le versioni del cluster non sono interi semplici. Per trovare la versione corrente del cluster, usa l'DescribeClusteroperazione o il comando AWS CLI describe-cluster. Una versione di esempio è KTVPDKIKX0DER.

```
aws kafka update-cluster-kafka-version --cluster-arn ClusterArn --current-
version Current-Cluster-Version --target-kafka-version TargetVersion
```

L'output del comando precedente è simile al JSON seguente.

```
{

    "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/exampleClusterName/
abcdefab-1234-abcd-5678-cdef0123ab01-2",
    "ClusterOperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef"
}
```

3. Per ottenere il risultato dell'`update-cluster-kafka-version`operazione, esegui il comando seguente, sostituendo *ClusterOperationArn* con l'ARN ottenuto nell'output del `update-cluster-kafka-version` comando.

```
aws kafka describe-cluster-operation --cluster-operation-arn ClusterOperationArn
```

L'output di questo comando `describe-cluster-operation` è simile all'esempio JSON seguente.

```
{
    "ClusterOperationInfo": {
        "ClientRequestId": "62cd41d2-1206-4ebf-85a8-dbb2ba0fe259",
```

```
        "ClusterArn": "arn:aws:kafka:us-east-1:012345678012:cluster/
exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2",
        "CreationTime": "2021-03-11T20:34:59.648000+00:00",
        "OperationArn": "arn:aws:kafka:us-east-1:012345678012:cluster-
operation/exampleClusterName/abcdefab-1234-abcd-5678-cdef0123ab01-2/0123abcd-
abcd-4f7f-1234-9876543210ef",
        "OperationState": "UPDATE_IN_PROGRESS",
        "OperationSteps": [
            {
                "StepInfo": {
                    "StepStatus": "IN_PROGRESS"
                },
                "StepName": "INITIALIZE_UPDATE"
            },
            {
                "StepInfo": {
                    "StepStatus": "PENDING"
                },
                "StepName": "UPDATE_APACHE_KAFKA_BINARIES"
            },
            {
                "StepInfo": {
                    "StepStatus": "PENDING"
                },
                "StepName": "FINALIZE_UPDATE"
            }
        ],
        "OperationType": "UPDATE_CLUSTER_KAFKA_VERSION",
        "SourceClusterInfo": {
            "KafkaVersion": "2.4.1"
        },
        "TargetClusterInfo": {
            "KafkaVersion": "2.6.1"
        }
    }
}
```

Se il valore di `OperationState` è `UPDATE_IN_PROGRESS`, attendi qualche minuto, quindi esegui nuovamente il comando `describe-cluster-operation`. Al termine dell'operazione, il valore di `OperationState` diventa `UPDATE_COMPLETE`. Poiché il tempo necessario ad Amazon MSK per completare l'operazione varia, potrebbe essere necessario eseguire ripetutamente il controllo fino al completamento dell'operazione.

Aggiornamento della versione di Apache Kafka utilizzando l'API

1. Richiama l'GetCompatibleKafkaVersionsoperazione per ottenere un elenco delle versioni di Apache Kafka a cui è possibile aggiornare il cluster.
2. Richiama l'UpdateClusterKafkaVersionoperazione per aggiornare il cluster a una delle versioni compatibili di Apache Kafka.

Le migliori pratiche per gli aggiornamenti delle versioni

Per garantire la continuità del client durante l'aggiornamento progressivo eseguito come parte del processo di aggiornamento della versione di Kafka, rivedi la configurazione dei tuoi client e gli argomenti di Apache Kafka come segue:

- Imposta il fattore di replica dell'argomento (RF) su un valore minimo di per i cluster Two-AZ e un valore minimo di 2 per i cluster Three-AZ. 3 Un valore RF di 2 può portare a partizioni offline durante l'applicazione delle patch.
- Imposta il numero minimo di repliche in sincronia (miniSR) su un valore massimo di per garantire che il set di repliche delle partizioni possa `RF - 1` tollerare che una replica sia offline o poco replicata.
- Configura i client per utilizzare più stringhe di connessione del broker. La presenza di più broker nella stringa di connessione di un client consente il failover se un broker specifico che supporta l'I/O del client inizia a ricevere le patch. Per informazioni su come ottenere una stringa di connessione con più broker, consulta Ottenere i broker bootstrap per un cluster Amazon MSK.
- Ti consigliamo di aggiornare i client che si connettono alla versione consigliata o superiore per beneficiare delle funzionalità disponibili nella nuova versione. Gli upgrade dei client non sono soggetti alle date di fine del ciclo di vita (EOL) della versione Kafka del cluster MSK e non è necessario che vengano completati entro la data di fine del ciclo di vita. Apache Kafka fornisce una politica di compatibilità dei client bidirezionale che consente ai client più vecchi di lavorare con cluster più recenti e viceversa.
- È probabile che i client Kafka che utilizzano le versioni 3.x.x abbiano le seguenti impostazioni predefinite: e. `acks=all` `enable.idempotence=true` `acks=all`è diverso dall'impostazione predefinita precedente di `acks=1` e offre una maggiore durabilità assicurando che tutte le repliche sincronizzate riconoscano la richiesta di produzione. Analogamente, l'impostazione predefinita per `enable.idempotence` era precedente. `false` La modifica all'`enable.idempotence=true`impostazione predefinita riduce la probabilità di messaggi

duplicati. Queste modifiche sono considerate impostazioni di best practice e possono introdurre una piccola quantità di latenza aggiuntiva che rientra nei normali parametri di prestazione.

- Utilizzate la versione consigliata di Kafka per creare nuovi cluster MSK. L'utilizzo della versione consigliata di Kafka consente di sfruttare le funzionalità più recenti di Kafka e MSK.

Risoluzione dei problemi relativi al cluster Amazon MSK

Le seguenti informazioni consentono di semplificare la risoluzione dei problemi che si potrebbero verificare con il cluster Amazon MSK. Puoi anche pubblicare il problema in AWS re:Post.

Argomenti

La sostituzione del volume causa la saturazione del disco a causa del sovraccarico della replica

In caso di guasto hardware non pianificato del volume, Amazon MSK può sostituire il volume con una nuova istanza. Kafka ripopola il nuovo volume replicando le partizioni di altri broker del cluster. Una volta che le partizioni sono state replicate e recuperate, sono idonee per l'iscrizione alla leadership e all'In-Sync Replica (ISR).

Problema

In un broker che si sta riprendendo dalla sostituzione dei volumi, alcune partizioni di dimensioni diverse potrebbero tornare online prima di altre. Ciò può essere problematico in quanto tali partizioni possono servire il traffico proveniente dallo stesso broker che sta ancora recuperando (replicando) altre partizioni. Questo traffico di replica a volte può saturare i limiti di throughput del volume sottostanti, che nel caso predefinito sono 250 MiB al secondo. Quando si verifica questa saturazione, tutte le partizioni già interessate ne risentono, con conseguente latenza all'interno del cluster per tutti i broker che condividono ISR con quelle partizioni interessate (non solo le partizioni leader dovute agli ack remoti). `acks=all` Questo problema è più comune nei cluster più grandi che hanno un numero maggiore di partizioni di dimensioni variabili.

Raccomandazione

- Per migliorare la postura I/O della replica, assicuratevi che siano state adottate le migliori impostazioni dei thread.
- Per ridurre la probabilità di saturazione del volume sottostante, abilita lo storage fornito con un throughput più elevato. Un valore di throughput minimo di 500 MIB/s è consigliato per i casi di replica ad alto throughput, ma il valore effettivo necessario varia in base alla velocità effettiva e al caso d'uso. Assegnazione della velocità di trasmissione effettiva dell'archiviazione.
- Per ridurre al minimo la pressione di replica, `num.replica.fetchers` abbassare al valore predefinito di2.

Gruppo di consumatori bloccato nello stato `PreparingRebalance`

Se uno o più gruppi di consumatori sono bloccati in uno stato di ribilanciamento perpetuo, la causa potrebbe essere il problema KAFKA-9752 di Apache Kafka, che riguarda le versioni 2.3.1 e 2.4.1 di Apache Kafka.

Per risolvere questo problema, ti consigliamo di aggiornare il cluster alla versione Versione di correzione dei bug Amazon MSK 2.4.1.1, che contiene una correzione per questo problema. Per informazioni sull'aggiornamento di un cluster esistente alla versione 2.4.1.1 di correzione dei bug di Amazon MSK, consulta la pagina Aggiornamento della versione di Apache Kafka.

Le soluzioni alternative per risolvere questo problema senza aggiornare il cluster alla versione di correzione del bug Amazon MSK 2.4.1.1 consistono nell'impostare i client Kafka in modo da utilizzare Protocollo di iscrizione statico oppure Identificazione e riavvio il nodo dei broker di coordinamento del gruppo di consumatori bloccato.

Implementazione del protocollo di iscrizione statico

Per implementare il protocollo di iscrizione statico nei client, procedi come indicato di seguito:

1. Imposta la proprietà `group.instance.id` della configurazione dei consumatori Kafka su una stringa statica che identifica il consumatore nel gruppo.
2. Assicurati che le altre istanze della configurazione siano aggiornate in modo da utilizzare la stringa statica.
3. Implementa le modifiche ai tuoi consumatori Kafka.

L'utilizzo del protocollo di iscrizione statico è più efficace se il timeout della sessione nella configurazione client è impostato su una durata che consenta al consumatore di ripristinare il sistema senza innescare prematuramente un ribilanciamento del gruppo di consumatori. Ad esempio, se l'applicazione consumatore può tollerare 5 minuti di indisponibilità, un valore ragionevole per il timeout della sessione sarebbe 4 minuti anziché il valore predefinito di 10 secondi.

> **Note**
>
> L'utilizzo del protocollo di iscrizione statico riduce solamente la probabilità di riscontrare questo problema. È possibile che questo problema si verifichi ancora anche quando si utilizza il protocollo di iscrizione statico.

Riavvio del nodo dei broker di coordinamento

Per riavviare il nodo dei broker di coordinamento, procedi come segue:

1. Identifica il coordinatore del gruppo utilizzando il comando `kafka-consumer-groups.sh`.
2. Riavvia il coordinatore del gruppo di consumatori bloccato utilizzando l'azione RebootBrokerAPI.

Errore nell'invio dei log del broker ad Amazon CloudWatch Logs

Quando provi a configurare il tuo cluster per inviare i log del broker ad Amazon CloudWatch Logs, potresti ottenere una delle due eccezioni.

Se viene restituita un'eccezione `InvalidInput.LengthOfCloudWatchResourcePolicyLimitExceeded`, riprova utilizzando i gruppi di log che iniziano con `/aws/vendedlogs/`. Per ulteriori informazioni, consulta la pagina Enabling Logging from Certain Amazon Web Services.

Se ricevi un'`InvalidInput.NumberOfCloudWatchResourcePoliciesLimitExceeded`eccezione, scegli una policy Amazon CloudWatch Logs esistente nel tuo account e aggiungi il seguente codice JSON.

```
{"Sid":"AWSLogDeliveryWrite","Effect":"Allow","Principal":
{"Service":"delivery.logs.amazonaws.com"},"Action":
["logs:CreateLogStream","logs:PutLogEvents"],"Resource":["*"]}
```

Se provi ad aggiungere il codice JSON sopra riportato a una policy esistente ma ricevi un errore che indica che hai raggiunto la lunghezza massima per la policy che hai scelto, prova ad aggiungere il codice JSON a un'altra delle tue politiche Amazon Logs. CloudWatch Dopo aver aggiunto il codice JSON a una policy esistente, prova ancora una volta a configurare la distribuzione dei log del broker ad Amazon Logs. CloudWatch

Nessun gruppo di sicurezza predefinito

Se cerchi di creare un cluster e ricevi un errore che indica che non esiste un gruppo di sicurezza predefinito, è possibile che il VPC che stai utilizzando sia stato condiviso con te. Chiedi all'amministratore di concedere l'autorizzazione per descrivere i gruppi di sicurezza in questo VPC e riprova. Per un esempio di una policy che consente questa operazione, consulta Amazon EC2: consente la gestione dei gruppi di sicurezza EC2 associati a uno specifico VPC, in modo programmatico e nella console .

I cluster sono bloccati nello stato CREATING

A volte la creazione del cluster può richiedere fino a 30 minuti. Attendi 30 minuti e controlla nuovamente lo stato del cluster.

Lo stato del cluster passa da CREATING a FAILED

Prova a creare nuovamente il cluster.

Lo stato del cluster è ACTIVE ma i produttori non possono inviare dati o i consumatori non possono ricevere dati

- Se la creazione del cluster va a buon fine (lo stato del cluster è `ACTIVE`), ma non è possibile inviare o ricevere dati, assicurati che le applicazioni produttore e consumatore dispongano dell'accesso al cluster. Per ulteriori informazioni, consulta le linee guida in the section called "Passaggio 3: creazione di un computer client".

- Se i produttori e i consumatori dispongono dell'accesso al cluster ma si verificano ancora problemi nella produzione e nel consumo di dati, è possibile che la causa sia riconducibile a KAFKA-7697, che influenza Apache Kafka versione 2.1.0 e può condurre a un deadlock in uno o più broker. Valuta la possibilità di eseguire la migrazione ad Apache Kafka 2.2.1, che non è influenzato da questo bug. Per informazioni sulla migrazione, consulta Migrazione.

AWS CLI non riconosce Amazon MSK

Se lo hai AWS CLI installato, ma non riconosce i comandi di Amazon MSK, esegui l'upgrade AWS CLI alla versione più recente. Per istruzioni dettagliate su come aggiornare AWS CLI, consulta

Installazione di AWS Command Line Interface. Per informazioni su come utilizzare per AWS CLI eseguire i comandi Amazon MSK, consultaCome funziona.

Le partizioni vengono messe offline o le repliche non sono sincronizzate

Questi sintomi possono essere causati da spazio su disco insufficiente. Per informazioni, consulta the section called “Lo spazio su disco è insufficiente”.

Lo spazio su disco è insufficiente

Vedere le best practice seguenti per gestire lo spazio su disco: the section called “Monitoraggio dello spazio su disco” e the section called “Regolazione dei parametri di conservazione dei dati”.

La memoria è insufficiente

Se il parametro `MemoryUsed` diventa alto o il parametro `MemoryFree` diventa basso, ciò non significa che ci sia un problema. Apache Kafka è progettato per utilizzare e gestire in maniera ottimale la massima quantità di memoria.

Il produttore ottiene NotLeaderForPartitionException

Questo è spesso un errore temporaneo. Impostare il parametro di configurazione `retries` del produttore su un valore più alto del valore corrente.

Partizioni sottoreplicate (URP) superiori a zero

Il parametro `UnderReplicatedPartitions` è importante da monitorare. In un cluster MSK integro, il valore di questo parametro è 0. Se è maggiore di zero, il motivo potrebbe essere uno dei seguenti.

- Se `UnderReplicatedPartitions` presenta un picco, è possibile che non sia stato effettuato il provisioning del cluster alle dimensioni corrette per gestire il traffico in entrata e in uscita. Per informazioni, consulta Best practice.
- Se `UnderReplicatedPartitions` è costantemente maggiore di 0 anche durante i periodi di traffico ridotto, è possibile che siano stati impostate ACL restrittive che non concedono ai broker l'accesso all'argomento. Per replicare le partizioni, i broker devono disporre dell'autorizzazione

per gli argomenti READ e DESCRIBE. L'argomento DESCRIBE viene concesso per impostazione predefinita con l'autorizzazione READ. Per informazioni sull'impostazione degli ACL, consulta Authorization and ACLs nella documentazione di Apache Kafka.

Il cluster ha argomenti chiamati __amazon_msk_canary e __amazon_msk_canary_state

Potresti notare che il tuo cluster MSK ha un argomento con il nome `__amazon_msk_canary` e un altro con il nome `__amazon_msk_canary_state`. Si tratta di argomenti interni che Amazon MSK crea e utilizza per i parametri diagnostici e di salute dei cluster. Questi argomenti sono di dimensioni trascurabili e non possono essere eliminati.

La replica delle partizioni ha esito negativo

Assicurati di non aver impostato le ACL su CLUSTER_ACTIONS.

Impossibile accedere al cluster con accesso pubblico attivato

Se il cluster ha attivato l'accesso pubblico, ma non riesci ancora ad accedervi da Internet, esegui i passaggi seguenti:

1. Assicurati che le regole in entrata del gruppo di sicurezza del cluster consentano il tuo indirizzo IP e la porta del cluster. Per un elenco dei numeri di porta del cluster, consulta la pagina the section called “Informazioni sulle porte”. Assicurati inoltre che le regole in uscita del gruppo di sicurezza consentano le comunicazioni in uscita. Per ulteriori informazioni sui gruppi di sicurezza e le rispettive regole in entrata e in uscita, consulta la pagina Security groups for your VPC nella Guida per l'utente di Amazon VPC.
2. Assicurati che il tuo indirizzo IP e la porta del cluster siano consentiti nelle regole in entrata dell'ACL della rete VPC del cluster. A differenza dei gruppi di sicurezza, le ACL di rete sono prive di stato. Ciò significa che è necessario configurarne le regole in entrata e in uscita. Nelle regole in uscita, consenti tutto il traffico (intervallo di porte: 0-65535) verso il tuo indirizzo IP. Per ulteriori informazioni, consulta la pagina Add and delete rules nella Guida per l'utente di Amazon VPC.
3. Assicurati di utilizzare la stringa bootstrap-brokers ad accesso pubblico per accedere al cluster. Un cluster MSK con accesso pubblico attivato ha due diverse stringhe bootstrap-brokers, una per l'accesso pubblico e una per l'accesso dall'interno di AWS. Per ulteriori informazioni, consulta the section called “Ottenere i broker bootstrap utilizzando il AWS Management Console”.

Impossibile accedere al cluster dall'interno AWS: problemi di rete

Se disponi di un'applicazione Apache Kafka che non è in grado di comunicare correttamente con un cluster MSK, inizia eseguendo il seguente test di connettività.

1. Utilizzare uno dei metodi descritti in the section called “Recupero dei broker di bootstrap” per ottenere gli indirizzi dei broker bootstrap.
2. Nel seguente comando, sostituire *bootstrap-broker* con uno degli indirizzi dei broker ottenuti nella fase precedente. Sostituire *numero-porta* con 9094 se il cluster è configurato per utilizzare l'autenticazione TLS. Se il cluster non utilizza l'autenticazione TLS, sostituisci il *numero-porta* con 9092. Eseguire il comando dal computer client.

   ```
   telnet bootstrap-broker port-number
   ```

   Dove il numero di porta è:

   - 9094 se il cluster è configurato per utilizzare l'autenticazione TLS.
   - 9092 Se il cluster non utilizza l'autenticazione TLS.
   - È necessario un numero di porta diverso se l'accesso pubblico è abilitato.

   Eseguire il comando dal computer client.
3. Ripetere il comando precedente per tutti i broker bootstrap.

Se la macchina client è in grado di accedere ai broker, significa che non ci sono problemi di connettività. In questo caso, eseguire il comando seguente per verificare se il client Apache Kafka è configurato correttamente. Per ottenere *broker-bootstrap*, utilizzare uno dei metodi descritti in the section called “Recupero dei broker di bootstrap”. Sostituire l'*argomento* con il nome dell'argomento.

```
<path-to-your-kafka-installation>/bin/kafka-console-producer.sh --broker-
list bootstrap-brokers --producer.config client.properties --topic topic
```

Se il comando precedente va a buon fine, significa che il client è configurato correttamente. Se non è ancora possibile produrre e consumare da un'applicazione, eseguire il debug del problema a livello di applicazione.

Se la macchina client non è in grado di accedere ai broker, consulta le seguenti sottosezioni per una guida basata sulla configurazione del computer client.

Client Amazon EC2 e cluster MSK nello stesso VPC

Se il computer client si trova nello stesso VPC del cluster MSK, assicurati che il gruppo di sicurezza del cluster disponga di una regola in entrata che accetta il traffico dal gruppo di sicurezza del computer client. Per informazioni sull'impostazione di queste regole, consulta Regole del gruppo di sicurezza. Per un esempio di come accedere a un cluster da un'istanza Amazon EC2 che si trova nello stesso VPC del cluster, consulta la pagina Nozioni di base.

Client Amazon EC2 e cluster MSK in VPC diversi

Se il computer client e il cluster si trovano in due VPC diversi, verificare quanto segue:

- I due VPC sono in peering.
- Lo stato della connessione peering è attivo.
- Le tabelle di routing dei due VPC sono configurate correttamente.

Per informazioni sul peering VPC, consulta Utilizzo di connessioni peering VPC.

Client locale

Nel caso di un client locale configurato per connettersi al cluster MSK utilizzando, accertatevi di quanto segue: AWS VPN

- Lo stato della connessione VPN è UP. Per informazioni su come verificare lo stato della connessione VPN, consulta How do I check the current status of my VPN tunnel?.
- La tabella di routing del VPC del cluster contiene la route per un CIDR locale la cui destinazione ha il formato `Virtual private gateway(vgw-xxxxxxxx)`.
- Il gruppo di sicurezza del cluster MSK consente il traffico sulla porta 2181, sulla porta 9092 (se il cluster accetta traffico non crittografato) e sulla porta 9094 (se il cluster accetta traffico crittografato TLS).

Per ulteriori indicazioni AWS VPN sulla risoluzione dei problemi, consulta Troubleshooting Client VPN.

AWS Direct Connect

Se il client utilizza AWS Direct Connect, consulta Risoluzione dei problemi AWS Direct Connect.

Se le linee guida per risoluzione dei problemi precedenti non consentono di risolvere il problema, assicurarsi che il traffico di rete non sia bloccato da un firewall. Per ulteriori operazioni di debug, utilizza strumenti come `tcpdump` e `Wireshark` per analizzare il traffico e assicurarti che raggiunga il cluster MSK.

Autenticazione non riuscita: troppe connessioni

L'errore `Failed authentication ... Too many connects` indica che un broker si sta proteggendo perché uno o più client IAM stanno tentando di connettersi ad esso a una velocità aggressiva. Per aiutare i broker ad accettare nuove connessioni IAM a una velocità più elevata, puoi aumentare il parametro di configurazione `reconnect.backoff.ms`.

Per ulteriori informazioni sui limiti di velocità per le nuove connessioni per broker, consulta la pagina Quota di Amazon MSK.

MSK Serverless: la creazione del cluster ha esito negativo

Se si tenta di creare un cluster MSK Serverless e il flusso di lavoro ha esito negativo, è possibile che non si disponga dell'autorizzazione per creare un endpoint VPC. Verifica che l'amministratore ti abbia concesso l'autorizzazione a creare un endpoint VPC consentendo l'operazione `ec2:CreateVpcEndpoint`.

Per un elenco completo delle autorizzazioni necessarie per eseguire tutte le operazioni di Amazon MSK, consulta la pagina AWS politica gestita: AmazonMSK FullAccess.

Best practice

In questo argomento vengono illustrate alcune best practice da seguire quando si utilizza Amazon MSK.

Dimensionamento corretto del cluster: numero di partizioni per broker

Nella tabella seguente viene illustrato il numero consigliato di partizioni (incluse le repliche leader e follower) per broker.

Dimensioni del broker	Numero consigliato di partizioni (incluse le repliche leader e follower) per broker
`kafka.t3.small`	300
`kafka.m5.large` o `kafka.m5.xlarge`	1000
`kafka.m5.2xlarge`	2000
`kafka.m5.4xlarge` , `kafka.m5.8xlarge` , `kafka.m5.12xlarge` , `kafka.m5.16xlarge` oppure `kafka.m5.24xlarge`	4000
`kafka.m7g.large` o `kafka.m7g.xlarge`	1000
`kafka.m7g.2xlarge`	2000
`kafka.m7g.4xlarge` , `kafka.m7g.8xlarge` `kafka.m7g.12xlarge` , o `kafka.m7g.16xlarge`	4000

Se il numero di partizioni per broker supera il valore consigliato e il cluster si sovraccarica, è possibile che venga impedito di eseguire le seguenti operazioni:

- Aggiornamento della configurazione del cluster

- Aggiorna il cluster a un broker di dimensioni inferiori
- Associa un AWS Secrets Manager segreto a un cluster con autenticazione SASL/SCRAM

Un numero elevato di partizioni può inoltre comportare la mancanza delle metriche di Kafka CloudWatch su e sullo scraping di Prometheus.

Per informazioni sulla scelta del numero di partizioni, consulta Apache Kafka Supports 200K Partitions Per Cluster. Ti consigliamo inoltre di eseguire i tuoi test per determinare la dimensione giusta per i tuoi broker. Per ulteriori informazioni sulle diverse dimensioni dei broker, consultathe section called “Dimensioni dei broker”.

Dimensionamento corretto del cluster: numero di broker per cluster

Per determinare il numero corretto di broker per il cluster MSK e comprendere i costi, consulta il foglio di calcolo MSK Sizing and Pricing. Questo foglio di calcolo fornisce una stima delle dimensioni di un cluster MSK e dei costi associati di Amazon MSK rispetto a un cluster Apache Kafka simile basato su EC2 e autogestito. Per ulteriori informazioni sui parametri di input nel foglio di calcolo, passare il mouse sulle descrizioni dei parametri. Le stime fornite da questo foglio sono conservative e forniscono un punto di partenza per un nuovo cluster. Le prestazioni, le dimensioni e i costi del cluster dipendono dal caso d'uso e consigliamo di verificarli con test ad hoc.

Per capire in che modo l'infrastruttura sottostante influisce sulle prestazioni di Apache Kafka, consulta le migliori pratiche per il corretto dimensionamento dei cluster Apache Kafka per ottimizzare prestazioni e costi nel Big Data Blog. AWS Il post del blog fornisce informazioni su come dimensionare i cluster per soddisfare i requisiti di velocità di trasmissione effettiva, disponibilità e latenza. Fornisce inoltre risposte a domande quali quando è necessario aumentare o ridurre la capacità e indicazioni su come verificare continuamente le dimensioni dei cluster di produzione.

Ottimizza la velocità effettiva del cluster per istanze m5.4xl, m7g.4xl o più grandi

Quando si utilizzano istanze m5.4xl, m7g.4xl o più grandi, è possibile ottimizzare il throughput del cluster ottimizzando le configurazioni num.io.threads e num.network.threads.

Il valore num.io.threads è il numero di thread utilizzati da un broker per l'elaborazione delle richieste. L'aggiunta di più thread, fino al numero di core CPU supportati per la dimensione dell'istanza, può contribuire a migliorare il throughput del cluster.

Il valore num.network.threads è il numero di thread utilizzati dal broker per ricevere tutte le richieste in arrivo e restituire le risposte. I thread di rete inseriscono le richieste in entrata in una coda di richieste per l'elaborazione da parte di io.threads. L'impostazione di num.network.threads sulla metà del numero di core CPU supportati per la dimensione dell'istanza consente l'utilizzo completo della nuova dimensione dell'istanza.

> **Important**
>
> Non aumentare num.network.threads senza prima aumentare num.io.threads, in quanto ciò può causare una congestione legata alla saturazione della coda.

Impostazioni consigliate

Dimensioni istanza	Valore consigliato per num.io.threads	Valore consigliato per num.network.threads
m5.4xl	16	8
m5.8xl	32	16
m5.12xl	48	24
m5.16xl	64	32
m5.24xl	96	48
m7g.4xlarge	16	8
m7g.8xlarge	32	16
m7g.12xlarge	48	24
m7g.16xlarge	64	32

Usa l'ultima versione di Kafka per evitare problemi di mancata corrispondenza tra gli ID degli argomenti AdminClient

L'ID di un argomento viene perso (Errore: non corrisponde all'ID dell'argomento per la partizione) quando si utilizza una versione di Kafka AdminClient precedente alla 2.8.0 con il flag per aumentare o riassegnare le partizioni degli argomenti `--zookeeper` per un cluster utilizzando la versione di Kafka 2.8.0 o successiva. Nota che il flag `--zookeeper` è obsoleto in Kafka 2.5 ed è stato rimosso a partire da Kafka 3.0. Consulta la pagina Upgrading to 2.5.0 from any version 0.8.x through 2.4.x.

Per evitare la mancata corrispondenza degli ID degli argomenti, utilizza una versione del client Kafka 2.8.0 o successiva per le operazioni di amministrazione di Kafka. In alternativa, i client 2.5 e versioni successive possono utilizzare il flag `--bootstrap-servers` al posto del flag `--zookeeper`.

Creazione di cluster a disponibilità elevata

Utilizza i seguenti consigli in modo che il tuo cluster MSK possa essere altamente disponibile durante un aggiornamento (ad esempio quando aggiorni le dimensioni del broker o la versione di Apache Kafka, ad esempio) o quando Amazon MSK sostituisce un broker.

- Configura un cluster con tre zone di disponibilità.
- Assicurati che il fattore di replica (RF) sia almeno 3. Tieni presente che un valore di RF pari a 1 può portare a partizioni offline durante un aggiornamento in sequenza, mentre un RF pari a 2 può causare la perdita di dati.
- Impostare le repliche in sinc minime (minISR) su al massimo RF - 1. Un minISR uguale a RF può impedire la produzione nel cluster durante un aggiornamento in sequenza. Un minISR di 2 consente di rendere disponibili argomenti replicati a tre vie quando una replica è offline.
- Assicurati che le stringhe di connessione del client includano almeno un broker per ogni zona di disponibilità. La presenza di più broker nella stringa di connessione di un client consente il failover quando un broker specifico è offline a seguito di un aggiornamento. Per informazioni su come ottenere una stringa di connessione con più broker, consulta the section called “Recupero dei broker di bootstrap”.

Monitoraggio dell'utilizzo della CPU

Amazon MSK consiglia vivamente di mantenere l'utilizzo totale della CPU per i broker (definito come `CPU User + CPU System`) al di sotto del 60%. Quando hai a disposizione almeno il 40% della

CPU totale del cluster, Apache Kafka può ridistribuire il carico della CPU tra i broker del cluster, se necessario. Ad esempio, ciò si rende necessario quando Amazon MSK rileva e ripristina un errore del broker; in questo caso, Amazon MSK esegue la manutenzione automatica, ad esempio l'applicazione di patch. Un altro esempio è quando un utente richiede una modifica delle dimensioni del broker o un aggiornamento di versione; in questi due casi, Amazon MSK implementa flussi di lavoro continui che portano offline un broker alla volta. Quando i broker con partizioni leader vanno offline, Apache Kafka riassegna la leadership delle partizioni per ridistribuire il lavoro agli altri broker del cluster. Seguendo questa best practice, puoi assicurarti di avere abbastanza margine nella CPU nel cluster per tollerare eventi operativi come questi.

Puoi utilizzare Amazon CloudWatch Metric Math per creare una metrica composita. `CPU User + CPU System` Imposta un allarme che si attiva quando il parametro composito raggiunge un utilizzo medio della CPU del 60%. Quando viene attivato questo allarme, dimensiona il cluster utilizzando una delle seguenti opzioni:

- Opzione 1 (consigliata): aggiorna la dimensione del broker alla dimensione successiva più grande. Ad esempio, se la dimensione corrente è`kafka.m5.large`, aggiorna il cluster da utilizzare`kafka.m5.xlarge`. Tieni presente che quando aggiorni le dimensioni dei broker nel cluster, Amazon MSK disconnette i broker in modo continuativo e riassegna temporaneamente la leadership delle partizioni ad altri broker. Un aggiornamento delle dimensioni richiede in genere 10-15 minuti per broker.
- Opzione 2: se ci sono argomenti in cui tutti i messaggi sono stati acquisiti da produttori che utilizzano scritture ininterrotte (in altre parole, i messaggi non sono codificati e l'ordinamento non è importante per i consumatori), espandi il cluster aggiungendo altri broker. Inoltre, aggiungi partizioni agli argomenti esistenti con la velocità di trasmissione effettiva più elevata. Successivamente, utilizza `kafka-topics.sh --describe` per assicurarti che le partizioni appena aggiunte vengano assegnate ai nuovi broker. Il vantaggio principale di questa opzione rispetto alla precedente è la possibilità di gestire risorse e costi in modo più granulare. Inoltre, è possibile utilizzare questa opzione se il carico della CPU supera in modo significativo il 60%, poiché questa forma di dimensionamento in genere non comporta un aumento del carico per i broker esistenti.
- Opzione 3: espandi il cluster aggiungendo broker, quindi riassegna le partizioni esistenti utilizzando lo strumento di riassegnazione delle partizioni denominato `kafka-reassign-partitions.sh`. Tuttavia, se utilizzi questa opzione, il cluster dovrà spendere risorse per replicare i dati da broker a broker dopo la riassegnazione delle partizioni. Rispetto alle due opzioni precedenti, questa opzione può inizialmente aumentare in modo significativo il carico sul cluster. Di conseguenza, Amazon MSK sconsiglia di utilizzare questa opzione quando l'utilizzo della CPU è superiore al 70%, perché

la replica causa un carico aggiuntivo della CPU e del traffico di rete. Amazon MSK consiglia di utilizzare questa opzione solo se le due opzioni precedenti non sono percorribili.

Altre raccomandazioni:

- Monitora l'utilizzo totale della CPU per broker come proxy per la distribuzione del carico. Se i broker hanno un utilizzo della CPU costantemente irregolare, potrebbe essere un segno che il carico non è distribuito uniformemente all'interno del cluster. Amazon MSK consiglia di utilizzare Cruise Control per gestire in modo continuo la distribuzione del carico tramite l'assegnazione delle partizioni.
- Monitora la latenza di produzione e utilizzo. La latenza di produzione e utilizzo può aumentare linearmente con l'utilizzo della CPU.
- Intervallo di scrape JMX: se si abilita il monitoraggio aperto con la funzionalità Prometheus, si consiglia di utilizzare un intervallo di scrape di 60 secondi o superiore (scrape_interval: 60s) per la configurazione dell'host Prometheus (prometheus.yml). La riduzione dell'intervallo di scrape può comportare un utilizzo elevato della CPU sul cluster.

Monitoraggio dello spazio su disco

Per evitare di esaurire lo spazio su disco per i messaggi, crea un CloudWatch allarme che controlli la metrica. `KafkaDataLogsDiskUsed` Quando il valore di questo parametro raggiunge o supera l'85%, esegui una o più delle seguenti operazioni:

- Utilizza the section called “Scalabilità automatica”. Puoi anche aumentare manualmente lo spazio di archiviazione del broker come descritto nella sezione the section called “Dimensionamento manuale”.
- Riduci il periodo di conservazione dei messaggi o la dimensione del log. Per informazioni su come eseguire queste operazioni, consulta the section called “Regolazione dei parametri di conservazione dei dati”.
- Elimina argomenti non utilizzati.

Per informazioni su come configurare e utilizzare gli allarmi, consulta Using Amazon CloudWatch Alarms. Per un elenco completo di parametri di Amazon MSK, consulta la sezione Monitoraggio di un cluster.

Regolazione dei parametri di conservazione dei dati

Il consumo di messaggi non li rimuove dal log. Per liberare regolarmente spazio su disco, puoi specificare in modo esplicito un periodo di conservazione, ovvero il periodo di permanenza dei messaggi nel log. Puoi inoltre specificare una dimensione del log di conservazione. Quando viene raggiunto il periodo di conservazione o la dimensione del log di conservazione, Apache Kafka inizia a rimuovere i segmenti inattivi dal log.

Per specificare una policy di conservazione a livello di cluster, imposta uno o più dei seguenti parametri: `log.retention.hours`, `log.retention.minutes`, `log.retention.ms` o `log.retention.bytes`. Per ulteriori informazioni, consulta the section called “Configurazioni personalizzate di ”.

Puoi anche specificare i parametri di conservazione a livello di argomento:

- Per specificare un periodo di conservazione per argomento, utilizza il comando seguente.

```
kafka-configs.sh --bootstrap-server $bs --alter --entity-type topics --entity-
name TopicName --add-config retention.ms=DesiredRetentionTimePeriod
```

- Per specificare una dimensione del log di conservazione per argomento, utilizza il comando seguente.

```
kafka-configs.sh --bootstrap-server $bs --alter --entity-type topics --entity-
name TopicName --add-config retention.bytes=DesiredRetentionLogSize
```

I parametri di conservazione specificati a livello di argomento hanno la precedenza sui parametri a livello di cluster.

Accelerazione del ripristino dei log dopo un arresto non corretto

Dopo un arresto non corretto, un broker può impiegare del tempo per riavviarsi poiché esegue il ripristino dei log. Per impostazione predefinita, Kafka utilizza solo un thread per directory di log per eseguire questo ripristino. Ad esempio, se si dispone di migliaia di partizioni, il completamento del ripristino dei log può richiedere ore. Per velocizzare il ripristino dei log, si consiglia di aumentare il numero di thread utilizzando la proprietà di configurazione `num.recovery.threads.per.data.dir`. È possibile impostarlo sul numero di core CPU.

Monitoraggio della memoria di Apache Kafka

Ti consigliamo di monitorare la memoria utilizzata da Apache Kafka. In caso contrario, il cluster potrebbe diventare non disponibile.

Per determinare la quantità di memoria utilizzata da Apache Kafka, puoi monitorare il parametro `HeapMemoryAfterGC`. `HeapMemoryAfterGC` è la percentuale di memoria heap totale utilizzata dopo la rimozione di oggetti inutili (garbage collection). Ti consigliamo di creare un CloudWatch allarme che agisca quando un `HeapMemoryAfterGC` aumento supera il 60%.

Le operazioni che è possibile eseguire per ridurre l'utilizzo della memoria variano. Dipendono dal modo in cui si configura Apache Kafka. Ad esempio, se si utilizza la consegna transazionale dei messaggi, è possibile ridurre il valore `transactional.id.expiration.ms` nella configurazione di Apache Kafka da `604800000` ms a `86400000` ms (da 7 giorni a 1 giorno). Ciò riduce l'ingombro di memoria di ciascuna transazione.

Non aggiungere broker non MSK

Per i cluster ZooKeeper basati, se si utilizzano ZooKeeper i comandi Apache per aggiungere broker, questi broker non vengono aggiunti al cluster MSK e Apache ZooKeeper conterrà informazioni errate sul cluster. Ciò potrebbe comportare la perdita di dati. Per le operazioni cluster supportate, consulta Come funziona.

Abilitazione della crittografia dei dati in transito

Per informazioni sulla crittografia dei dati in transito e su come abilitarla, consulta the section called “Crittografia in transito”.

Riassegnazione delle partizioni

Per spostare le partizioni in broker diversi sullo stesso cluster, puoi utilizzare lo strumento di riassegnazione delle partizioni denominato `kafka-reassign-partitions.sh`. Ad esempio, dopo aver aggiunto nuovi broker per espandere un cluster o aver spostato le partizioni per rimuovere i broker, è possibile ribilanciare il cluster riassegnando le partizioni ai nuovi broker. Per informazioni su come aggiungere broker a un cluster, consulta the section called “Espansione di un cluster”. Per informazioni su come rimuovere i broker da un cluster, vedere. the section called “Rimuovi un

broker" Per informazioni sullo strumento di riassegnazione delle partizioni, consulta la sezione relativa all'espansione del cluster nella documentazione di Apache Kafka.

Cronologia del documento Guida per gli sviluppatori di Amazon MSK

Nella tabella seguente sono descritte le modifiche importanti apportate alla Guida per gli sviluppatori di Amazon MSK.

Ultimo aggiornamento della documentazione: 25 giugno 2024

Modifica	Descrizione	Data
Aggiunta la funzionalità Graviton upgrade in place.	È possibile aggiornare le dimensioni del cluster broker da M5 o T3 a M7g o da M7g a M5.	2024-6-25
3.4.0 Annunciata la data di fine del supporto.	La data di fine del supporto per la versione 3.4.0 di Apache Kafka è il 17 giugno 2025.	2024-6-24
Aggiunta la funzionalità di rimozione del broker.	È possibile ridurre la capacità di storage e di elaborazi one del cluster assegnato rimuovendo set di broker, senza alcun impatto sulla disponibilità, rischi di durabilit à dei dati o interruzione delle applicazioni di streaming dei dati.	2024-5-16
`WriteDataIdempoten tly` aggiunto a AWSMSKRep licatorExecutionRole	WriteDataIdempotently è stata aggiunta l'autoriz zazione alla AWSMSKRep licatorExecutionRole policy per supportare la replica dei dati tra cluster MSK.	2024-5-16

Modifica	Descrizione	Data
Broker Graviton M7g rilasciati in Brasile e Bahrain.	Amazon MSK ora supporta la disponibilità nelle regioni Sud America (sa-east-1, San Paolo) e Medio Oriente (me-south-1, Bahrain) dei broker M7g che utilizzano processori Graviton (processori personalizzati basati su ARM creati da Amazon Web Services). AWS	2024-2-07
Rilascia i broker Graviton M7g nella regione cinese	Amazon MSK ora supporta la disponibilità nella regione cinese dei broker M7g che utilizzano processori AWS Graviton (processori personalizzati basati su ARM creati da Amazon Web Services).	11/01/2024
Politica di supporto della versione di Amazon MSK Kafka	È stata aggiunta una spiegazione della politica di supporto della versione di Kafka supportata da Amazon MSK. Per ulteriori informazioni, consulta le versioni di Apache Kafka.	2023-12-08

Modifica	Descrizione	Data
Nuova politica dei ruoli di esecuzione del servizio per supportare Amazon MSK Replicator.	Amazon MSK ha aggiunto una nuova `AWSMSKReplicatorExecutionRole` policy per supportare Amazon MSK Replicator. Per ulteriori informazioni, consulta l'argomento relativo alle policy gestite da AWS : `AWSMSKReplicatorExecutionRole`.	2023-12-06
Supporto M7g Graviton	Amazon MSK ora supporta i broker M7g che utilizzano processori AWS Graviton (processori personalizzati basati su ARM creati da Amazon Web Services).	2023-11-27
Replicatore Amazon MSK	Il replicatore Amazon MSK è una nuova funzionalità che puoi utilizzare per replicare i dati tra cluster Amazon MSK. Amazon MSK Replicator include un aggiornamento della policy di FullAccess AmazonMSK. Per ulteriori informazioni, consulta l'argomento relativo alle policy gestite da AWS : `AmazonMSKFullAccess`.	28/09/2023

Modifica	Descrizione	Data
Sono state aggiornate le best practice IAM.	Guida aggiornata per l'allinea mento alle best practice IAM. Per ulteriori informazioni, consulta Best practice per la sicurezza in IAM.	-08
Aggiornamenti del ruolo collegato ai servizi per supportare la connettività privata multi-VPC	Amazon MSK ora include aggiornamenti dei ruoli AWSServiceRoleForKafka collegati ai servizi per gestire le interfacce di rete e gli endpoint VPC nel tuo account che rendono i broker di cluster accessibili ai clienti nel tuo VPC. Amazon MSK utilizza le autorizzazioni per `DescribeVpcEndpoints` , `ModifyVpcEndpoint` e `DeleteVpcEndpoints` . Per ulteriori informazioni, consulta Utilizzo di ruoli collegati ai servizi per Amazon MSK.	1-08
Supporto per Apache Kafka 2.7.2	Amazon MSK ora supporta Apache Kafka versione 2.7.2. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2021-12-21
Supporto per Apache Kafka 2.6.3	Amazon MSK ora supporta Apache Kafka versione 2.6.3. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2021-12-21

Modifica	Descrizione	Data
Versione preliminare di MSK Serverless	MSK Serverless è una nuova funzionalità che è possibile utilizzare per creare cluster serverless. Per ulteriori informazioni, consulta MSK Serverless.	2021-11-29
Supporto per Apache Kafka 2.8.1	Amazon MSK ora supporta Apache Kafka versione 2.8.1. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2021-09-30
MSK Connect	MSK Connect è una nuova funzionalità che è possibile utilizzare per creare e gestire i connettori Apache Kafka. Per ulteriori informazioni, consulta MSK Connect.	2021-09-16
Supporto per Apache Kafka 2.7.1	Amazon MSK ora supporta Apache Kafka versione 2.7.1. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2021-05-25
Supporto per Apache Kafka 2.8.0	Amazon MSK ora supporta Apache Kafka versione 2.8.0. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2021-04-28

Modifica	Descrizione	Data
Supporto per Apache Kafka 2.6.2	Amazon MSK ora supporta Apache Kafka versione 2.6.2. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2021-04-28
Supporto per l'aggiornamento del tipo di broker	È ora possibile modificare il tipo di broker per un cluster esistente. Per ulteriori informazioni, consulta Aggiornamento delle dimensioni del broker.	2021-01-21
Supporto per Apache Kafka 2.6.1	Amazon MSK ora supporta Apache Kafka versione 2.6.1. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2021-01-19
Supporto per Apache Kafka 2.7.0	Amazon MSK ora supporta Apache Kafka versione 2.7.0. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2020-12-29

Modifica	Descrizione	Data
Indisponibilità di nuovi cluster con Apache Kafka versione 1.1.1	Non è più possibile creare un nuovo cluster Amazon MSK con la versione 1.1.1 di Apache Kafka. Tuttavia, se disponi di cluster MSK esistenti che eseguono Apache Kafka versione 1.1.1, puoi continuare a utilizzare tutte le funzionalità attualmen te supportate su tali cluster esistenti. Per ulteriori informazi oni, consulta Versioni di Apache Kafka.	2020-11-24
Parametri relativi al ritardo dei consumatori	Ora Amazon MSK offre parametri che permettono di monitorare il ritardo dei consumatori. Per ulteriori informazioni, consulta Monitoraggio di un cluster Amazon MSK.	2020-11-23
Supporto per Cruise Control	Amazon MSK ora supporta il LinkedIn Cruise Control. Per ulteriori informazioni, consulta Utilizzo LinkedIn del Cruise Control per Apache Kafka con Amazon MSK.	2020-11-17
Supporto per Apache Kafka 2.6.0	Amazon MSK ora supporta Apache Kafka versione 2.6.0. Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	2020-10-21

Modifica	Descrizione	Data
Supporto per Apache Kafka 2.5.1	Amazon MSK ora supporta Apache Kafka versione 2.5.1. Con la versione 2.5.1 di Apache Kafka, Amazon MSK supporta la crittografia in transito tra client ed endpoint. ZooKeeper Per ulteriori informazioni, consulta Versioni di Apache Kafka supportate.	30/09/2020
Espansione automatica dell'applicazione	Puoi configurare Streaming gestito da Amazon per Apache Kafka per espandere automaticamente l'archiviazione del tuo cluster in risposta a un incremento dell'utilizzo. Per ulteriori informazioni, consulta Scalabilità automatica.	2020-09-30
Supporto per la sicurezza di nome utente e password	Amazon MSK ora supporta l'accesso ai cluster utilizzando nome utente e password. Amazon MSK archivia le credenziali in AWS Secrets Manager. Per ulteriori informazioni, consulta Autenticazione SASL/SCRAM.	2020-09-17
Supporto per l'aggiornamento della versione Apache Kafka di un cluster Amazon MSK	Puoi aggiornare la versione di Apache Kafka di un cluster MSK esistente.	2020-05-28

Modifica	Descrizione	Data
Supporto per nodi broker T3.Small	Ora Amazon MSK supporta la creazione di cluster con broker di tipo T3.small di Amazon EC2.	2020-04-08
Supporto per Apache Kafka 2.4.1.	Amazon MSK ora supporta Apache Kafka versione 2.4.1.	2020-04-02
Supporto per i log del broker di streaming	Amazon MSK ora può trasmettere i log dei broker a CloudWatch Logs, Amazon S3 e Amazon Data Firehose. Firehose può, a sua volta, consegnare questi log alle destinazioni supportate, come Service. OpenSearch	2020-02-25
Supporto per Apache Kafka 2.3.1.	Amazon MSK ora supporta Apache Kafka versione 2.3.1.	19-12-2019
Monitoraggio aperto	Amazon MSK ora supporta il monitoraggio aperto con Prometheus.	04-12-2019
Supporto per Apache Kafka 2.2.1.	Amazon MSK ora supporta Apache Kafka versione 2.2.1.	31-07-2019
Disponibilità generale	Le nuove caratteristiche includono il supporto di tagging, l'autenticazione, la crittografia TLS, le configura zioni e la possibilità di aggiornare lo storage broker.	30-05-2019
Supporto per Apache Kafka 2.1.0.	Amazon MSK ora supporta Apache Kafka versione 2.1.0.	05-02-2019

AWS Glossario

Per la AWS terminologia più recente, consultate il AWS glossario nella sezione Reference.Glossario AWS

Le traduzioni sono generate tramite traduzione automatica. In caso di conflitto tra il contenuto di una traduzione e la versione originale in Inglese, quest'ultima prevarrà.