Cloudera Stream Processing 2.0.1

Cloudera Stream Processing Deployment

Date published: 2019-11-08
Date modified: 2019-11-08

CLOUDERA

https://docs.cloudera.com/

Legal Notice

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

CSP Deployment Overview

It is helpful to understand an overview of the Cloudera Stream Processing (CSP) deployment process before you begin. Deploying CSP preparing your environment and cluster, securing your cluster, then setting up a secure Apache Kafka cluster, and then adding and configuring your Schema Registry, Streams Replication Manager, and Streams Messaging Manager.

Important:

CSP 2.0.x installs Streaming components on CDH clusters. For information about installing Streaming components on CDP Private Cloud Base, see *Runtime documentation* for your software version. For information about installing Streaming components in CDP Public Cloud, see *CDF for Data Hub documentation* about Streams Messaging clusters.

Related Information

Runtime documentation

CDF for Data Hub documentation

Prepare your environment

Before you begin your CSP deployment, ensure that you have installed the required JDK and the database you want to use as the relational data store for Schema Registry and Streams Messaging Manager.

Install the JDK

You should install JDK 8 on each machine on which you will install CSP, if your version of Cloudera Manager is not already configured with JDK 8.

Procedure

1. Download JDK from the appropriate website.
2. Run the installation command appropriate for your operating system:

   For RHEL/CentOS/SLES:

   ```
   yum install java-1.8.0-openjdk-devel
   ```

   For Ubuntu:

   ```
   apt-get install openjdk-8-jdk
   ```

Related Information

Open JDK Download

Oracle JDK Download

Installing Databases

If you are installing Schema Registry or Streams Messaging Manager (SMM), you require a relational data store to store metadata. You can use either MySQL, Postgres, or Oracle. These topics describe how to install MySQL and Postgres, and how create a database for Schema Registry and SMM.

Note:

- You should install either Postgres or MySQL. Both are note required. MySQL is recommended.
- For a full list of supported databases and versions, see the *CSP Support Matrix*.

Related Information

CSP Support Matrix

Installing MySQL

Learn how to install MySQL databases for CSP components

Procedure

1. Log in to the node on which you want to install the MySQL metastore.
2. Install MySQL and the MySQL community server, and start the MySQL service:

```
yum localinstall \
https://dev.mysql.com/get/mysql57-community-release-el7-8.noarch.rpm
yum install mysql-community-server
systemctl start mysqld.service
```

3. Obtain the randomly generated MySQL root password.

```
grep 'A temporary password is generated for root@localhost' \
/var/log/mysqld.log |tail -1
```

4. Reset the MySQL root password. Enter the following command. You are prompted for the password you obtained in the previous step. MySQL then asks you to change the password.

```
/usr/bin/mysql_secure_installation
```

Configuring Metadata Stores in MySQL

Once you install a MySQL database, you must configure it for Schema Registry and SMM.

Procedure

1. Launch the MySQL monitor:

```
mysql -u root -p
```

2. Create the database for the Schema Registry and the SMM metastore:

```

create database registry;
create database streamsmsgmgr;
```

3. Create Schema Registry and SMM user accounts, replacing the final IDENTIFIED BY string with your password:

```

CREATE USER 'registry'@'%' IDENTIFIED BY 'R12$%34qw';
CREATE USER 'streamsmsgmgr'@'%' IDENTIFIED BY 'R12$%34qw';
```

4. Assign privileges to the user account:

```

GRANT ALL PRIVILEGES ON registry.* TO 'registry'@'%' WITH GRANT OPTION ;
GRANT ALL PRIVILEGES ON streamsmsgmgr.* TO 'streamsmsgmgr'@'%' WITH GRANT
OPTION ;
```

   If you cannot grant all privileges, grant the following privileges that SMM and Schema Registry require at a minimum:

   - CREATE/ALTER/DROP TABLE
   - CREATE/ALTER/DROP INDEX
   - CREATE/ALTER/DROP SEQUENCE

- CREATE/ALTER/DROP PROCEDURE

For example:

```
grant create session to streamsmsgmgr;
grant create table to streamsmsgmgr;
grant create sequence to streamsmsgmgr;
```

5. Commit the operation:

```
commit;
```

Install Postgres

About this task

If you have already installed a MySQL database, you may skip these steps.

Procedure

1. Install Red Hat Package Manager (RPM) according to the requirements of your operating system:

```
yum install https://yum.postgresql.org/9.6/redhat/rhel-7-x86_64/repoview/
postgresql96.html
```

2. Install Postgres version 9.5 or later:

```
yum install postgresql96-server postgresql96-contrib postgresql96
```

3. Initialize the database:

   - For CentOS 7, use the following syntax:

```
/usr/pgsql-9.6/bin/postgresql96-setup initdb
```

   - For CentOS 6, use the following syntax:

```
sudo service postgresql initdb
```

4. Start Postgres.

For example, if you are using CentOS 7, use the following syntax:

```

systemctl enable postgresql-9.6.service
systemctl start postgresql-9.6.service
```

5. Verify that you can log in:

```

sudo su postgres
psql
```

Configure Postgres to Allow Remote Connections

Configuring Postgres to allow remote connections involves editing the pg_hba.conf and postgresql.conf files and restarting Postgres.

Before you begin

It is critical that you configure Postgres to allow remote connections before you deploy a cluster. If you do not perform these steps in advance of installing your cluster, the installation fails.

Procedure

1. Open /var/lib/pgsql/9.6/data/pg_hba.conf and update to the following

```

# "local" is for Unix domain socket connections only
local all all trust


# IPv4 local connections:
host all all 0.0.0.0/0 trust

# IPv6 local connections:
host all all ::/0 trust
```

2. Open /var/lib//pgsql/9.6/data/postgresql.conf and update to the following:

```
listen_addresses = '*'
```

3. Restart Postgres:

```
systemctl stop postgresql-9.6.service
systemctl start postgresql-9.6.service
```

Configure Metadata Stores in Postgres

If you have already installed MySQL and configured the metadata store using MySQL, you do not need to configure additional metadata stores in Postgres.

Procedure

1. Log in to Postgres:

```

sudo su postgres
psql
```

2. For the Schema Registry metadata store, create a database called registry with the password registry:

```

create database registry;
CREATE USER registry WITH PASSWORD 'registry';
GRANT ALL PRIVILEGES ON DATABASE "registry" to registry;
```

3. For the SMM metadata store, create a database called streamsmsgmgr with the password streamsmsgmgr:

```

create database streamsmsgmgr;
CREATE USER streamsmsgmgr WITH PASSWORD 'streamsmsgmgr';
GRANT ALL PRIVILEGES ON DATABASE "streamsmsgmgr" to streamsmsgmgr;
```

   If you cannot grant all privileges, grant the following privileges that SMM and Schema Registry require at a minimum:

   - CREATE/ALTER/DROP TABLE
   - CREATE/ALTER/DROP INDEX
   - CREATE/ALTER/DROP SEQUENCE
   - CREATE/ALTER/DROP PROCEDURE

   For example:

```
grant create session to streamsmsgmgr;
grant create table to streamsmsgmgr;
grant create sequence to streamsmsgmgr;
```

Prepare your cluster

Prepare your CSP cluster by installing or upgrading Cloudera Manager, installing or upgrading CDH, and installing Kafka either as part of the CDH 6.3.0 or later parcel or using the CDK 4.1 or later parcel for CDH 5.13.x or later.

Install or upgrade Cloudera Manager

Ensure that you have installed or upgraded to Cloudera Manager 6.3.x or later.

Related Information

Cloudera Manager installation documentation

Cloudera Manager upgrade documentation

Install CDH

Install or upgrade to the version of CDH you want to run. Supported CDH version are CDH 5.13 and later, and CDH 6.3.0 and later.

About this task:

When you install CDH you will install Apache Kafka in one of two ways:

- If you install CDH 6.3.0 or later, Apache Kafka is included in your CDH installation.
- If you are using CDH 5.13.0 or later, you must install Apache Kafka using the CDK 4.1 parcel.

Related Information

CDH 6.3.x installation documentation

CDH 6.3.x upgrade documentation

CDH 5.16.x installation documentation

CDH 5.16.x upgrade documentation

Configure Apache Kafka for SMM

After you have installed and configured Apache Kafka, you must set one configuration parameter to enable Kafka and SMM to communicate.

Before you begin

- You have installed Apache Kafka either by installing CDH 6.3.0 or later, or by using the CDK 4.1 parcel to install Apache Kafka onto CDH 5.13.0 or later.
- You have configured Kafka according to your use case needs. Use the information available in *Apache Kafka Administration* to configure Kafka as needed.

Procedure

1. Select Kafka from your cluster drop-down, and then select the Configuration tab.
2. Ensure that the Enable Producer Metrics check box is selected.

Related Information

Configuring High Availability and Consistency for Apache Kafka

Configuring Apache Kafka for Performance and Resource Management

Apache Kafka Administration

Secure your cluster

Authenticate your cluster by enabling Kerberos on your cluster, install Sentry, and enabling Kerberos on Sentry.

Note:

The following security configuration steps are strongly recommended for your production environment, but are not required.

Enable Kerberos on your cluster

Procedure

1. From the Cloudera Manager Administration drop-down, select Security.
2. Select Enable Kerberos and walk through the additional prompts

Related Information

Enabling Kerberos Authentication for CDH

Install Sentry

Sentry is an RPC server that stores authorization metadata in an underlying relational database and provides RPC interfaces to retrieve and manipulate privileges. It supports secure access to services using Kerberos.

Before you begin:

- If you are using CDH 6.3.0 or later, you must install HDFS.
- If you are using CDH 5.13.0 or later, you must install either HDFS or Isilon.

Related Information

Before you install Sentry

Installing and Upgrading the Sentry Service

Configuring the Sentry Service

Enable Auto-TLS for your cluster

After you have enabled Kerberos, you must configure TLS. To configure TLS for Cloudera Manager and CSP Services, follow the instructions provided in the *Cloudera Manager documentation*.

Related Information

Cloudera Manager Documentation for Configuring Auto-TLS

Set up a secure Apache Kafka cluster

Securing an Apache Kafka cluster includes enabling Kerberos, TLS, and Apache SentryYou can enable Kafka security authentication using Kerberos. Ensure that you use the SASL_SSL Kafka protocol. Find complete details for securing Kafka in the *Kafka security documentation*.

Related Information

Kafka security documentation for CDH 6.3.0

Kafka security documentation for CDK 4.1 on CDH 5.13.0 or later

Install the Parcel and CSD files

To install the CSP services, copy the parcels and CSD files into the respective directories, restart cloudera managers, and then download, distribute, and activate the CSP service parcels.

About this task

Install the services in this order:

- Schema Registry
- SRM
- SMM

Before you begin

- You have obtained the CSD, parcel and .sha files from the Cloudera Archive using your authentication credentials (username/password) provided as part of the subscription fulfillment process. You can use your license file to obtain the authentication credentials by generating them on the CDP Private Cloud Base download page after logging in and selecting Download Now.
- You have installed Apache Kafka.

Procedure

1. Copy the parcel and .sha files into the parcel directory. By default, the parcel repository is /opt/cloudera/parcel-repo, located on the server where Cloudera Manager is running.
2. Copy the CSD files into the CSD directory. By default, this is located at /opt/cloudera/csd, located on the server where Cloudera Manager is running.
3. Change the parcel, .sha, and CSD files ownership and permissions. Enter:

   ```
   # For parcels
   chown cloudera-scm:cloudera-scm /opt/cloudera/parcel-repo/*
   chmod 644 /opt/cloudera/parcel-repo/*
   # For CSDs
   chown cloudera-scm:cloudera-scm /opt/cloudera/csd/*
   chmod 644 /opt/cloudera/csd/*
   ```
4. Restart the Cloudera Manager Server.

   For RHEL 7

   ```
   sudo systemctl restart cloudera-scm-server
   ```

   For RHEL 6

   ```
   sudo service cloudera-scm-server restart
   ```
5. Restart the Cloudera Management Service. From the Cloudera Manager Status tab, click the Restart Icon or select Restart from the Cloudera Management Service drop-down.
6. From the Cloudera Manager navigation bar, click Hosts then Parcels, and click the Configuration button.
7. Review the parcel list and verify that the Schema Registry, SRM, and SMM parcels are available. If not, click the Check for New Parcels button. The new parcels should display Available for Download on the Parcels page.
8. From the Parcels page, download, distribute, and activate the Schema Registry, SRM, and SMM parcels.

Related Information

Download locations

Add and Configure Schema Registry

Procedure

1. From the Cloudera Manager Home page, select the drop-down to the right of your cluster, and select Add a Service and select Schema Registry. You may install one service at a time.
2. If you updated the JDK in the *Prepare your environment* section of this document, ensure that you also update the Java Home Path Override field in the Cloudera Manager configuration.
3. Specify the Jar Storage Type.

   Your options are:

   - Local
   - HDFS

   For example:

   ```
   schema.registry.jar.storage.type=local
   ```
4. Configure the Registry Storage properties, based on the database you created to use as the Schema Registry metadata store.

   ```
   Schema Registry Storage Connector URL / schema.registry.storage.connecto
   r.connectURI

   Schema Registry Storage Connector User / schema.registry.storage.connector
   .user
   Schema Registry Storage Connector Password / schema.registry.storage.conne
   ctor.password
   ```
5. Ensure that Schema Registry Storage Connector URL has the fully qualified host name for the database installation location, the connector URL, and the default port for the database selected. For example:

   ```
   MYSQL example:
   jdbc:mysql://FQDN_MYSQL:3306/registry

   Postgres example:
   jdbc:postgresql://FQDN_POSTGRES:5432/registry

   Oracle example:
   jdbc:oracle:thin:@FQDN_ORACLE:3306/registry
   ```
6. Complete the rest of the Cloudera Manager wizard steps to complete the Schema Registry installation and configuration.

Add and Configure SRM

Add and Configure SRM

Learn how to add SRM to your cluster.

About this task

Streams Replication Manager is comprised of two roles:

- Streams Replication Manager Driver role: This role is responsible for connecting to the specified clusters and performing replication between them. The driver can be installed on one or more hosts.
- Streams Replication Manager Service role: This role consist of a REST API and a Kafka Streams application to aggregate and expose cluster, topic and consumer group metrics. The service can be installed on one host only.

You can install Streams Replication Manager independent of the clusters that replication is happening between.

The following steps walk you through the process of adding SRM to your cluster. The configuration examples on this page are simple examples that are meant to demonstrate the type of information that you have to enter. For comprehensive deployment and configuration examples, see Deployment recommendations and Configuration examples.

Note: Do not confuse Streams Replication Manager, which is a service managed by Cloudera Manager, with the Streams Replication Manager Service role, which is a role within Streams Replication Manager.

Before you begin

If Kafka is configured to use Sentry authorization, make sure that streamsrepmgr is added to the Kafka Super users property.

Procedure

1. From the Cloudera Manager Home page, select the drop-down to the right of your cluster, and select Add a Service.
2. Select Streams Replication Manager from the list of services and click Continue.
3. Assign role instances to hosts:

   Note: In certain cases role names on this page are incorrectly displayed and may become truncated. The Streams Replication Manager Driver role is the role displayed on the left, while the Streams Replication Manager Service role is the role displayed on the right.

   a) Click the field below Streams Replication Manager Driver to display a dialog box containing a list of hosts.
   b) Select 1 or more hosts that the Streams Replication Manager Driver should be assigned to and Click Ok.
   c) Click the field below Streams Replication Manager Service to display a dialog box containing a list of hosts.
   d) Select 1 host that the Streams Replication Manager Service should be assigned to and Click Ok.
4. Click Continue.
5. Specify cluster aliases:

   a) Find the Streams Replication Manager Cluster alias. property.
   b) Add a comma delimited list of cluster aliases. For example:

   ```
   primary, secondary
   ```

   Cluster aliases are arbitrary names defined by the user. Aliases specified here are used in other configuration properties and with the srm-control tool to refer to the clusters added for replication.
6. Specify cluster connection information:

   a) Find the streams.replication.manager's replication configs property.
   b) Click the add button and add new lines for each cluster alias you have specified in the Streams Replication Manager Cluster alias. property
   c) Add connection information for your clusters. For example:

   ```
   primary.bootstrap.servers=primary_host1:9092,primary_host2:9092,primary_
   host3:9092
   secondary.bootstrap.servers=secondary_host1:9092,secondary_host2:9092
   ,secondary_host3:9092
   ```

   Each cluster has to be added to a new line. If a cluster has multiple hosts, add them to the same line but delimit them with commas.

7. Add and enable replications:
   a) Find the streams.replication.manager's replication configs property.
   b) Click the add button and add new lines for each unique replication you want to add and enable.
   c) Add and enable your replications. For example:

   ```
   primary->secondary.enabled=true
   secondary->primary.enabled=true
   ```
8. Specify the Streams Replication Manager Service role target cluster:
   a) Find the streams replication manager service target cluster property.
   b) Add the target cluster alias. For example:

   ```
   secondary
   ```

   The target cluster is where the service gathers replication information from. Cloudera recommends that you deploy the service on every cluster and configure each instance of the service to target the cluster that it is running on.
9. Optional: Specify the Streams Replication Manager Driver role target clusters:
   a) Find the streams replication manager driver target cluster property.
   b) Add the cluster aliases that you want the driver role to target. For example:

   ```
   primary, secondary
   ```

   You can use the streams replication manager driver target cluster property to specify a subset of clusters that the driver should target or in other words write to. When this property is left empty (default) the driver will read from and write to all clusters added to SRMs configuration. When this property is set, the driver will collect data from all clusters, but will only write to the clusters specified in this property. This property becomes essential when you have an advanced deployment as it allows you to distribute replication workloads.
10. Configure properties not exposed in Cloudera Manager:

    SRM accepts a number of additional configuration properties that are not available in Cloudera Manager. Depending on your requirements you may need to configure these properties as well. You can find a comprehensive list of these properties in Configuration Properties Reference.
    a) Find the streams.replication.manager's replication configs property.
    b) Click the add button and add new lines for each additional property you want to configure.
    c) Add configuration properties. For example:

    ```
    replication.factor=3
    ```
11. Depending on your requirement, review and configure other properties available on this page.
12. Click Continue and wait until Streams Replication Manager is started.
13. Click Continue then click Finish.

Results

- Replicating data to or from the specified clusters is now possible.
- The SRM service REST API Swagger UI is available at one of the following addresses:

  -
    ```
    http://<srm-service-host>:<srm-service-port/swagger
    ```

  -
    ```
    https://<srm-service-host>:<srm-service-port/swagger
    ```

What to do next

- Enable Kerberos and TLS/SSL for SRM.

- Use the srm-control tool to kick off replication by adding topics or groups to the allowlist.

Related Information

Configuration Properties Reference

srm-control

Configuration Examples

Enable Kerberos on SRM

Procedure

1. From the Clusters drop-down, select Streams Replication Manager.
2. Click the Configuration tab.
3. Click Enable Kerberos Authentication to enable Kerberos authentication.

Enable TLS/SSL for SRM

Learn how to Enable TLS/SSL for SRM.

Procedure

1. In Cloudera Manager select Streams Replication Manager.
2. Go to Configuration.
3. Find and enable the Enable TLS/SSL for Streams Replication Manager Driver property.
4. Find and enable the Enable TLS/SSL for Streams Replication Manager Service property.

Note: If a third-party truststore is used, you have to manually configure and add truststore information via Cloudera Manager.

Add and Configure SMM

Configuring Cloudera Manager and Service Monitor

SMM requires high levels of Cloudera Manager and Service Monitor Service memory usage, and is important to configure them appropriately. For details see *Cloudera Manager tuning documentation* and *Service Monitor tuning documentation*.

Related Information

Cloudera Manager tuning documentation

Service Monitor tuning documentation

Obtain the Kafka service name

You will need the Kafka service name when you are configuring SMM. It is helpful to obtain it in advance.

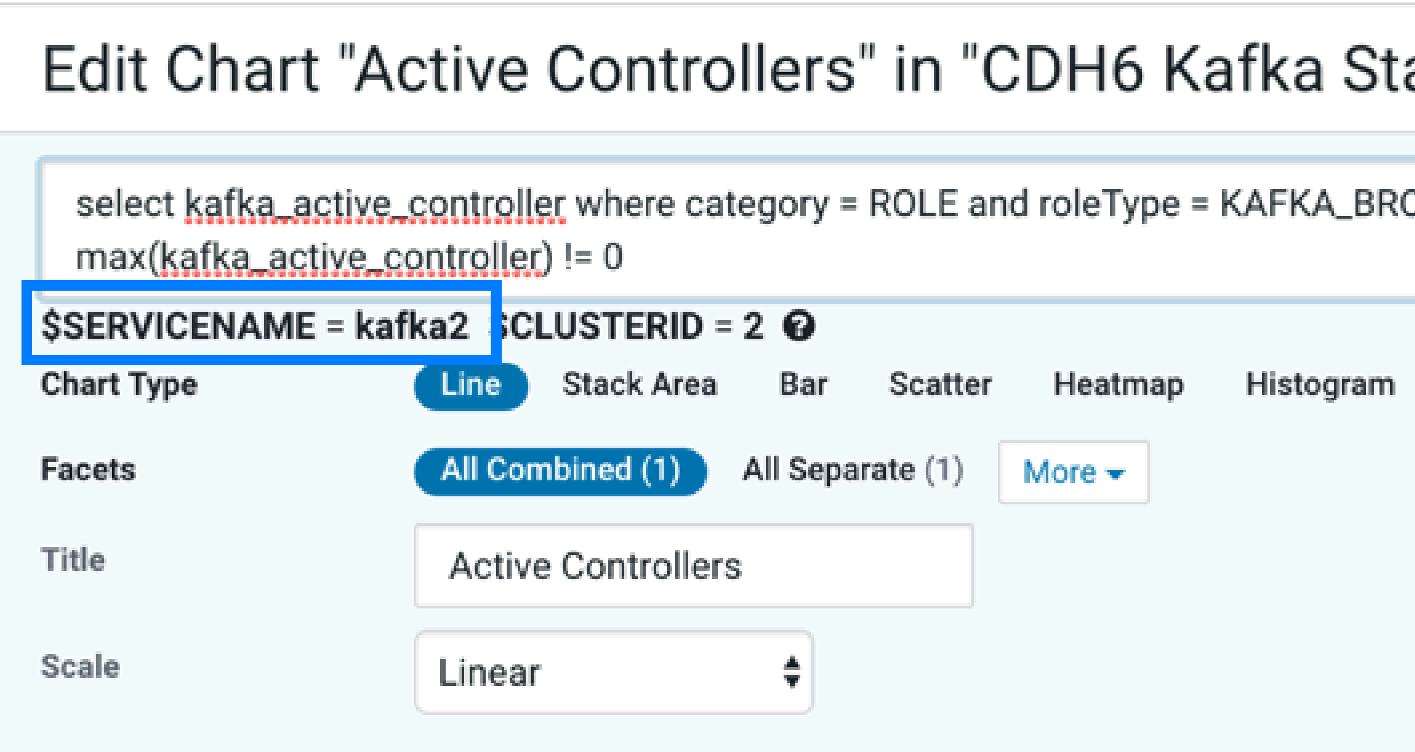
Procedure

1. In Cloudera Manager, go to your Kafka service.
2. Select any available chart, and select Open in Chart Builder from the configuration icon drop-down.
3. Find $SERVICENAME= near the top of the display.

Results

This value is the Kafka service name that you can use for SMM configuration.

Example



Edit Chart "Active Controllers" in "CDH6 Kafka Sta

```
select kafka_active_controller where category = ROLE and roleType = KAFKA_BRO
max(kafka_active_controller) != 0
```

$SERVICENAME = kafka2 $CLUSTERID = 2

Chart Type: Line

Facets: All Combined (1)

Title: Active Controllers

Scale: Linear

Add and configure the SMM Node

There are a number of required configurations you must make before you complete the SMM installation.

About this task

It is recommended that you add SMM to a management node on your cluster. For performance optimization, it is not recommended that you add it to a node where a Kafka Broker is running.

You will be installing two SMM roles on the same node: SMM REST Server, and SMM-UI.

Before you begin

- You have installed NPM and the forever module on the SMM host.
- You have configured Cloudera Manager and Service Monitor memory.
- You have obtained the Kafka service name.

Procedure

1. From the Cloudera Manager Home page, select the drop-down to the right of your cluster, and select Add a Service and select Streams Messaging Manager. You may install one service at a time.
2. Select your service dependencies:
   - If you plan to enable Kerberos, select HDFS, Kafka, Sentry, and ZooKeeper.
   - If you do not plan to enable Kerberos, select Kafka, and ZooKeeper.
3. Select Streams Messaging Manager from your cluster, and then select the Configuration tab.

4. Configure the following required properties:

Property	Description

streams.messaging.manager.storage.connector.connectURI	Specify the SMM database connector URL. For example:
	`jdbc:mysql://FQDN_MYHSQL:3306/streamsmsgmgr,`
	`jdbc:postgresql://FQDN_POSTGRES:5432/streamsmsgmgr`

streams.messaging.manager.storage.connector.user	Specify the Streams Messaging Manager storage connector user.
streams.messaging.manager.storage.connector.password	Specify the Streams Messaging Manager storage connector password.
streams.messaging.manager.kafka.client.security.protocol	Specify the correct option based on the Kafka security protocol: • If Kafka is not secured, choose PLAINTEXT • If Kafka is Kerberos and SSL enabled, choose Inferred or SASL_SSL If Kafka has Kerberos but no SSL, choose SASL_PLAINTEXT If Kafka has no Kerberos but does have SSL enabled, choose SSL
cm.metrics.protocol	Set the value to https if TLS is enabled for Cloudera Manager. If TLS is not enabled, the value is http.
cm.metrics.host	Specify Cloudera Manager's FQDN host name.
cm.metrics.port	Specify Cloudera Manager's port number. By default, this is 7183 if TLS is configured and 7180 if it is not.
cm.metrics.password	Specify the Cloudera Manager password.
cm.metrics.service.name	Specify the Cloudera Manager Kafka service name. Provide the name of the Kafka service you obtained before you began SMM configuration. See *Obtain the Kafka service name* for more information.
Schema Registry API url	Specify the URL for the Schema Registry REST API. For example: `https://localhost:7790/api/v1`

5. Click Finish and once SMM is running restart Cloudera Manager if prompted to do so.

Enable Kerberos on SMM

Procedure

1. From the Clusters drop-down, select Streams Messaging Manager.
2. Click the Configuration tab.
3. Click Enable Kerberos Authentication.

Enable TLS for SMM

Learn how to enable TLS/SSL encryption for Streams Messaging Manager (SMM). You can enable the settings in Cloudera Manager according to the cluster configuration.

About this task

If Kerberos is enabled, then you must enable SSL for SMM. SMM UI fails to load if Kerberos is enabled and SSL is not enabled.

Also, if Kafka has Kerberos/SSL enabled, the same should be enabled for SMM.

Procedure

1. Go to Cloudera Manager.
2. Select Streams Messaging Manager cluster.
3. Click Configuration from the menu bar.
4. In the Search field, type TLS/SSL to show the SMM TLS/SSL properties.

   The security related properties appear.
5. Edit the security properties according to the cluster configuration.
6. Click Save Changes.

TLS/SSL settings for Streams Messaging Manager

To enable TLS/SSL settings for Streams Messaging Manager (SMM), you need to configure SMM server properties, SMM UI properties, and SMM Server’s Oracle TLS connection properties in Cloudera Manager according to the cluster configuration.

Properties	Description
SMM server properties	
Enable TLS/SSL for Streams Messaging Manager Rest Admin Server ssl.enable	Encrypt communication between clients and Streams Messaging Manager Rest Admin Server using Transport Layer Security (TLS) (formerly known as Secure Socket Layer (SSL)).
Streams Messaging Manager port (SSL) streams.messaging.manager.ssl.port	HTTPS port Streams Messaging Manager rest server runs on when SSL is enabled.
Streams Messaging Manager Admin Port (SSL) streams.messaging.manager.ssl.adminPort	HTTPS admin port Streams Messaging Manager rest server runs on when SSL is enabled.
SSL Keystore Type streams.messaging.manager.ssl.keyStoreType	The keystore type. Required if Streams Messaging Manager rest server's SSL is enabled. e.g. PKCS12 or JKS. If it is left empty then the keystore type will come from CM settings.
SSL TrustStore Type streams.messaging.manager.ssl.trustStoreType	The truststore type. Required if streams messaging manager's ssl is enabled. e.g. PKCS12 or JKS. If it is left empty then the keystore type will come from CM settings.
Streams Messaging Manager Rest Admin Server TLS/SSL Server JKS Keystore File Location streams.messaging.manager.ssl.keyStorePath	The path to the TLS/SSL keystore file containing the server certificate and private key used for TLS/SSL. Used when Streams Messaging Manager Rest Admin Server is acting as a TLS/SSL server.
Streams Messaging Manager Rest Admin Server TLS/SSL Server JKS Keystore File Password	The password for the Streams Messaging Manager Rest Admin Server keystore file.
Streams Messaging Manager Rest Admin Server TLS/SSL Server JKS Keystore Key Password	The password that protects the private key contained in the keystore used when Streams Messaging Manager Rest Admin Server is acting as a TLS/SSL server.
Streams Messaging Manager Rest Admin Server TLS/SSL Client Trust Store File streams.messaging.manager.ssl.trustStorePath	The location on disk of the trust store used to confirm the authenticity of TLS/SSL servers that Streams Messaging Manager Rest Admin Server might connect to. This is used when Streams Messaging Manager Rest Admin Server is the client in a TLS/SSL connection. This trust store must contain the certificate(s) used to sign the service(s) connected to. If this parameter is not provided, the default list of well-known certificate authorities is used instead.

Properties	Description
Streams Messaging Manager Rest Admin Server TLS/SSL Client Trust Store Password	The password for the Streams Messaging Manager Rest Admin Server TLS/SSL Certificate Trust Store File. This password is not required to access the trust store; this field can be left blank. This password provides optional integrity checking of the file. The contents of trust stores are certificates, and certificates are public information.
Cloudera Manager Metrics TrustStore Type cm.metrics.truststore.type	Cloudera Manager's truststore type. If it is left empty then the keystore type will come from CM settings. If it is left empty then the keystore type will come from CM settings.
SSL ValidateCerts streams.messaging.manager.ssl.validateCerts	Whether or not to validate TLS certificates before starting. If enabled, it will refuse to start with expired or otherwise invalid certificates.
SSL validatePeers streams.messaging.manager.ssl.validatePeers	Whether or not to validate TLS peer certificates.
SMM UI properties	
Enable TLS/SSL for Streams Messaging Manager UI Server streams.messaging.manager.ui.ssl.enable	Encrypt communication between clients and Streams Messaging Manager UI Server using Transport Layer Security (TLS) (formerly known as Secure Socket Layer (SSL)).
Streams Messaging Manager UI Server TLS/SSL Server Private Key File (PEM Format) streams.messaging.manager.ui.ssl.private.key.location	The path to the TLS/SSL file containing the private key used for TLS/SSL. Used when Streams Messaging Manager UI Server is acting as a TLS/SSL server. The certificate file must be in PEM format.
Streams Messaging Manager UI Server TLS/SSL Server Certificate File (PEM Format) streams.messaging.manager.ui.ssl.cert.location	The path to the TLS/SSL file containing the server certificate key used for TLS/SSL. Used when Streams Messaging Manager UI Server is acting as a TLS/SSL server. The certificate file must be in PEM format.
Streams Messaging Manager UI Server TLS/SSL Server CA Certificate (PEM Format) streams.messaging.manager.ui.ssl.ca.cert.location	The path to the TLS/SSL file containing the certificate of the certificate authority (CA) and any intermediate certificates used to sign the server certificate. Used when Streams Messaging Manager UI Server is acting as a TLS/SSL server. The certificate file must be in PEM format, and is usually created by concatenating all of the appropriate root and intermediate certificates.
Streams Messaging Manager UI Server TLS/SSL Private Key Password	The password for the private key in the Streams Messaging Manager UI Server TLS/SSL Server Certificate and Private Key file. If left blank, the private key is not protected by a password.
Streams Messaging Manager UI Server TLS/SSL Certificate Trust Store File streams.messaging.manager.ui.ssl.trust.store.location	The location on disk of the trust store, in .pem format, used to confirm the authenticity of TLS/SSL servers that Streams Messaging Manager UI Server might connect to. This is used when Streams Messaging Manager UI Server is the client in a TLS/SSL connection. This trust store must contain the certificate(s) used to sign the service(s) connected to. If this parameter is not provided, the default list of well-known certificate authorities is used instead.
SMM server's Oracle TLS connection properties	
Enable TLS with Oracle DB streams.messaging.manager.enable.TLS.Oracle	Enable TLS with Oracle as DB for Schema Registry.
Oracle.net.ssl_version streams.messaging.manager.oracle.net.ssl_version	Oracle net ssl version.
Oracle TLS javax.net.ssl.keyStore streams.messaging.manager.javax.net.ssl.keyStore	Path to keystore file if enabling TLS using Oracle DB.
Oracle TLS javax.net.ssl.keyStoreType streams.messaging.manager.javax.net.ssl.keyStoreType	KeyStoreType type if enabling TLS using Oracle DB.
Oracle TLS javax.net.ssl.keyStorePassword streams.messaging.manager.javax.net.ssl.keyStorePassword	KeyStorePassword if enabling TLS using Oracle DB.

Properties	Description
Oracle TLS javax.net.ssl.trustStore streams.messaging.manager.javax.net.ssl.trustStore	Required Path to truststore file if enabling TLS using Oracle DB.
Oracle TLS javax.net.ssl.trustStoreType streams.messaging.manager.javax.net.ssl.trustStoreType	Required Truststore type if enabling TLS using Oracle DB.
Oracle TLS javax.net.ssl.trustStorePassword streams.messaging.manager.javax.net.ssl.trustStorePassword	TrustStorePassword type if enabling TLS using Oracle DB.
Oracle TLS oracle.net.ssl_cipher_suites streams.messaging.manager.oracle.net.ssl_cipher_suites	net ssl cipher suites if enabling TLS using Oracle DB e.g. SSL_DH_DSS_WITH_DES_CBC_SHA.
Oracle TLS oracle.net.ssl_server_dn_match streams.messaging.manager.oracle.net.ssl_server_dn_match	ssl server domain name match if enabling TLS using Oracle DB.
Oracle TLS oracle.net.authentication_services streams.messaging.manager.oracle.net.authentication_services	Oracle net authentication service if enabling TLS using Oracle DB.

Enabling Sentry Authorization for SMM

You can use Sentry to configure and manage authorization for SMM.

Procedure

1. To enable Sentry Service admin permissions for the SMM user, add streamsmsgmgr to the sentry.service.admin.group field.
2. To add SMM to the list of allowed Sentry Service users, add streamsmsgmgr to the sentry.service.allow.connect field.
3. From the Kafka Configurations tab, click the + next to the super.users field and add streamsmsgmgr.
4. From the Kafka Configurations tab, search for Sentry Service and select the Sentry instance you want to use for Kafka authorization. For example: Sentry-1
5. From the SMM Configurations tab, search for Sentry Service and select the Sentry instance you want to use for SMM authorization. For example: Sentry-1
6. Restart Sentry.
7. Restart Kafka and redeploy the client configurations.
8. Restart SMM.

Note: SMM users need to be granted Sentry privileges to Kafka and the topics. The following is an example of Sentry permissions required for Kafka, for granting access to all the topics:

a. Create a Sentry role for SMM users.

```
$ kafka-sentry -cr -r smm_user_role
```

b. Assign the Sentry role to an SMM user group.

```
$ kafka-sentry -arg -r smm_user_role -g smm_user_group
```

c. Grant the required privilege to the Sentry role.

```
$ kafka-sentry -gpr -r smm_user_role -p "Host=*->Topic=*->actio
n=describe"
$ kafka-sentry -gpr -r smm_user_role -p "Host=*->Cluster=kafka-c
luster->action=describe"
$ kafka-sentry -gpr -r smm_user_role -p "Host=*->Consumergroup=*-
>action=describe"
```

Enable partition level metric collection

You must enable SMM to configure partition level metrics about Kafka topics.

Before you begin

You have obtained the Cloudera Manager Kafka service name. See *Obtain the Kafka service name* for information.

Procedure

1. In Cloudera Manager, select Hosts, and then Hosts Configuration.
2. To allow SMM to collect partition level metrics about Kafka topics, add the following to Cloudera Manager Agent Monitoring Advanced Configuration Snippet (Safety Valve):

   ```
   <serviceName>_broker_topic_partition_metrics_for_smm_enabled=true
   ```

   When you are configuring partition level metrics collection, your Kafka service name should always be in lower case letters.

   For example if KAFKA is the Kafka serviceName, enter the following:

   ```
   kafka_broker_topic_partition_metrics_for_smm_enabled=true
   ```
3. Restart Cloudera Manager if prompted to do so.

Configuring SMM to Monitor SRM Replications

If you plan to use SMM to monitor Kafka cluster replication, ensure that you have configured SMM to communicate with SRM.

Procedure

1. From the Clusters drop-down, select Streams Messaging Manager.
2. Click the Configuration tab.
3. Click the Configure Streams Replication Manager box.
4. Specify the SRM protocol, host, and port in the remaining fields.

   - The SRM host is a host where the Streams Replication Manager Service role is running.
   - The SRM service port is the port where Streams Replication Manager service is running. It is specified in the config streams.replication.manager.service.port field.
5. Restart Cloudera Manager if prompted to do so.