Cloudera Runtime 7.1.9

Streams Replication Manager Reference

Date published: 2019-08-22
Date modified: 2023-09-07

CLOUDERA

https://docs.cloudera.com/

Legal Notice

Contents

srm-control Options Reference

A collection of all options and their descriptions for the srm-control command line tool.

Table 1: Topics and groups subcommand properties

Options	Description
-h, --help	Shows the help message.
--source	Specifies the source cluster
--target	Specifies the target cluster
--config	Specifies the SRM configuration file to use.
--add	Specifies topics or groups to add to the allowlist
--remove	Specifies topics or groups to remove from the allowlist
--add-blacklist	Specifies topics or groups to add to the denylist
--remove-blacklist	Specifies topics or groups to remove from the denylist
--list	Lists current allowlist and denylist
--bootstrap-servers	Specifies the bootstraps servers
--producer-props	Specifies producer configuration properties
--consumer-props	Specifies consumer configuration properties
--props	Specifies client configuration properties

Table 2: Offsets subcommand properties

Option	Description
-h, --help	Shows the help message.
--source	Specifies the source cluster
--target	Specifies the target cluster
--config	Specifies the SRM configuration file to use.
--export	Export translated offsets
--group	Specifies the groups translated offsets should be exported for
--bootstrap-servers	Specifies the bootstraps servers
--props	Specifies client configuration properties

Configuration Properties Reference for Properties not Available in Cloudera Manager

A collection of SRM configuration properties not available in Cloudera Manager

The following table lists all SRM specific configuration properties that are not available directly for configuration via Cloudera Manager.

Property	Default Value	Description
checkpoints.topic.replication.factor	3	Replication factor used for internal checkpoints topics.
emit.checkpoints.enabled	true	Enables periodic emission of consumer offset information.
emit.heartbeats.enabled	true	Enables periodic emission of heartbeats.
emit.heartbeats.interval.seconds	5 (seconds)	The interval at which SRM emits heartbeats.
heartbeats.topic.replication.factor	3	Replication factor used for internal heartbeat topics.
offset-syncs.topic.replication.factor	3	Replication factor used for internal offset-syncs topics.
refresh.groups.enabled	true	Enables a periodical check for new consumer groups on source clusters.
refresh.topics.enabled	true	Enables a periodical check for new topics on source clusters.
replication.factor	2	Replication factor used for remote topics.
replication.policy.class	org.apache.kafka.connect.mirror.DefaultReplicationPolicy	Replication policy to use. Use this property to set a custom replication policy. If you want to enable prefixless replication, use the Enable Prefixless Replication property instead. Enable Prefixless Replication is found in Cloudera Manager SRM Configuration For more information, see *Enabling prefixless replication*.
sync.topic.configs.enabled	true	Enables the monitoring of the source cluster for configuration changes.

Related Information
Enabling prefixless replication

Kafka credentials property reference

A Kafka credential allows services and Cloudera tools to securely connect to Kafka clusters that are external to a particular Cloudera Manager instance. Review the Kafka credentials property reference to better understand what options you have when configuring and setting up a Kafka credential.

Name
Description

> Specifies the name of the Kafka credential. For SRM, the name you specify here is equivalent to the cluster alias. As a result, this name is used to refer to the cluster that the credential defines when configuring SRM properties and when using the srm-control tool. Cloudera recommends that you specify unique and easily identifiable names.

Boostrap servers
Description

> Specifies the addresses of the Kafka brokers. Add a comma-separated list of host and port pairs. The host can be specified as an IP address or an FQDN.

Example

> my-kafka-cluster-host-1.com:9093,   my-kafka-cluster-host-2.com:9093

JAAS secret [1-3]

Description

There are three JAAS Secret properties, numbered 1-3. All three properties can be used to specify sensitive JAAS configuration values. Values set in these properties can be referenced with placeholders in the JAAS configuration you add to the JAAS template property.

JAAS template

Description

Specifies the JAAS configuration that SRM uses when connecting to this cluster. The configuration you add to this property can contain placeholders. These placeholders can be used to hide sensitive information. The property accepts and resolves the following placeholders:

- ##JAAS_SECRET_1##

  Used to refer to the value set in the JAAS Secret 1 property.
- ##JAAS_SECRET_2##

  Used to refer to the value set in the JAAS Secret 2 property.
- ##JAAS_SECRET_3##

  Used to refer to the value set in the JAAS Secret 3 property.

Example

org.apache.kafka.common.security.plain.PlainLoginModule required   username="##JAAS_SECRET_1##" password="##JAAS_SECRET_2##";

In this example ##JAAS_SECRET_1## and ##JAAS_SECRET_2## are used to hide sensitive information. The values of the username and password are defined in the JAAS Secret 1 and JAAS Secret 2 properties.

If for example the JAAS Secret 1 and JAAS Secret 2 properties were set to "username" and "password", the resolved JAAS configuration would be the following:

```
org.apache.kafka.common.security.plain.PlainLoginModule required
  username="username" password="password";
```

Kerberos Service Name

Description

Specifies the name of the Kafka kerberos principal. Only required if the SASL Mechanism property is set to GSSAPI.

Key Password

Description

Specifies the password used to access the keys stored in the keystore configured in the Keystore Path property.

Keystore Password

Description

Specifies the password used to access the keystore configured in the Keystore Path property.

Keystore Path

Description

Specifies the path to the keystore containing the authorized client's certificates or keys

Keystore Type

Description

Specifies the type of the keystore configured in the Keystore Path property

SASL Mechanism

Description

Specifies the SASL mechanism used for authentication by the Kafka cluster that the client is connecting to.

Accepted values

- GSSAPI
- PLAIN
- SCRAM-SHA-256
- SCRAM-SHA-512
- OAUTHBEARER

Security Protocol

Description

Specifies the security protocol used by the Kafka cluster that the client is connecting to.

Accepted values

- PLAINTEXT
- SSL
- SASL_PLAINTEXT
- SASL_SSL

Truststore Password

Description

Specifies the password used to access the truststore specified in the Truststore Path property

Truststore Path

Description

Specifies the path to the truststore containing the certificates or keys belonging to the Kafka cluster that the client is connecting to.

Truststore Type

Description

Specifies the type of the truststore configured in the Truststore Path property.

SRM Service data traffic reference

The SRM Service generates data traffic when it collects replication related metrics. Data traffic is generated between the SRM Service and its target Kafka cluster. Additionally, if Remote Querying is set up, data traffic is also generated between different SRM Services that are remote to each other. Review these examples to gain a better understanding of how much data traffic you should expect in your deployment.

Data transferred between the SRM Service and Kafka

The SRM service generates data traffic between itself and its target Kafka cluster. This data traffic is generated as a result of the SRM Service collecting the raw metrics from the target Kafka cluster. In the case of:

- A single replication.

- A single replicated topic with 1000 partitions and 100 messages published per second.
- A single consumer group.

The amount of data traffic generated is approximately 280 kB per second.

Data transferred between SRM Services by Remote Querying

When Remote Querying is enabled, the SRM Service acting as the designated monitoring gateway generates additional data traffic. This data traffic is generated as a result of the gateway communicating with other, remote SRM Services and receiving data on remote replications. In the case of:

- A single replication flow.
- A single replicated topic with 1000 partitions.
- A single consumer group.
- One hour retention time set for the Streams application.

The maximum amount of data traffic generated by Remote Querying (the data returned by other SRM Services) is around 70 kB.