

Cloudera Streaming Analytics Operator 1.4.0

Cloudera Streaming Analytics - Kubernetes Operator 1.4.0 Release Notes

Date published: 2024-06-15

Date modified: 2025-10-24

The Cloudera logo is displayed in a bold, orange, sans-serif font. The word "CLOUDERA" is written in all caps, with a stylized 'E' that has a horizontal bar extending to the right.

<https://docs.cloudera.com/>

Legal Notice

© Cloudera Inc. 2025. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

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.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

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

Cloudera Streaming Analytics - Kubernetes Operator 1.4.0.....	4
What's new.....	4
Fixed issues and improvements.....	5
Improvements.....	5
Behavioral changes.....	5
Known issues and limitations.....	6
Unsupported features.....	7
 Deprecation notices in Cloudera Streaming Analytics - Kubernetes	
Operator.....	8
Deprecation Notices for Cloudera SQL Stream Builder.....	8
 Component support.....	9
 System requirements.....	10

Cloudera Streaming Analytics - Kubernetes Operator 1.4.0

Learn about the new features, improvements, known and fixed issues, limitations, and unsupported features in this release of Cloudera Streaming Analytics - Kubernetes Operator.

What's new

Learn about the new features and notable changes in this release of Cloudera Streaming Analytics - Kubernetes Operator.

Cloudera Streaming Analytics - Kubernetes Operator 1.4.0

To learn more about Cloudera Streaming Analytics - Kubernetes Operator and its typical deployment architecture, see [Overview](#) and for the installation instructions, see [Installation](#).

This release of the Cloudera Streaming Analytics - Kubernetes Operator is based on the Apache Flink Kubernetes Operator (Flink Operator) 1.11.0 and Apache Flink (Flink) 1.20. For more information about the supported versions, see the following upstream resources:

- [Flink Operator 1.11](#)
- [Flink 1.20](#)

What's new

Learn about the new features and notable changes in this release of Cloudera Streaming Analytics - Kubernetes Operator.

Cloudera Streaming Analytics - Kubernetes Operator 1.4.0

This release of Cloudera Streaming Analytics - Kubernetes Operator is based on the Apache Flink Kubernetes Operator (Flink Operator) 1.11 and Apache Flink (Flink) 1.20. For more information about the supported versions, see the following upstream resources:

- [Flink Operator 1.11](#)
- [Flink 1.20](#)

Python UDFs enabled by default

The Python User Defined Functions (UDFs) feature is enabled by default in Cloudera Streaming Analytics - Kubernetes Operator 1.4.0 deployments.

Removed v1 API from Cloudera SQL Stream Builder

The v1 REST API for Cloudera SQL Stream Builder is now removed from Cloudera Streaming Analytics.

Cloudera recommends migrating to the v2 API. For more information, see [Cloudera SQL Stream Builder REST API Reference](#).

Support FIPS Mode

Customers can configure and deploy the Cloudera Streaming Analytics - Kubernetes Operator in Federal Information Processing Standards (FIPS) mode.

This ensures that the deployment adheres to strict security and compliance requirements by utilizing FIPS-certified cryptographic modules, meeting the standards required for sensitive and regulated environments, such as the Veterans Administration.

Related Information

[Python UDFs in Cloudera SQL Stream Builder | Cloudera Streaming Analytics Documentation](#)

[Adjusting logging configuration in Advanced Settings | Cloudera Streaming Analytics Documentation](#)

Fixed issues and improvements

Learn about the fixed issues, improvements, and changes functionality in this release of Cloudera Streaming Analytics - Kubernetes Operator.

Fixed issues

CSA-5816: Cannot export an SSB project to GitHub that has secrets

Previously, exporting projects containing secrets to GitHub resulted in an empty dialog on Force Push. This issue is resolved by improving error handling when pushing sensitive data and updating data source naming.

CSA-5858: V2 API no longer allows executing jobs in session mode

Updated the API and connectors to no longer permit session mode execution.

CSA-5923: Stop with savepoint API hangs on Kubernetes

The "Stop with savepoint" API hung on Kubernetes due to a change in savepoint handling behavior. The API now verifies the upgradeSavepointPath to ensure a new savepoint is generated before stopping the job.

CSA-5930: Cannot stop session cluster even when only finished jobs are present on Kubernetes

Session clusters with only finished jobs could not be stopped due to the FlinkSessionJobs job not being automatically cleaned up. This is fixed by checking for globally terminal state and deleting the jobs so the session cluster may be terminated.

CSA-5933: Sampling fails on Kubernetes when incoming row kind is not INSERT

Sampling failed on Kubernetes for non-INSERT row kinds. This issue is fixed by formatting all incoming rows as INSERT.

Improvements

Learn about the fixed issues and improvements in this release of Cloudera Streaming Analytics - Kubernetes Operator.

Improvements

Learn about the change in certain functionality of Apache Flink and Cloudera SQL Stream Builder that has resulted in a change in behavior from the previously released version of Cloudera Streaming Analytics - Kubernetes Operator.

CSA-5940 - Add optimization to use single source for sampling in SSB

Summary: Optimized Flink SQL jobs to use a single source for sampling, reducing resource consumption and preventing job failures caused by duplicated source tasks.

CSA-5863 - Add option to specify truststore type for Kafka connections

Summary: Added support for specifying truststore and keystore types, such as JKS, PKCS12, or BCFKS, for Kafka connections in Global Config # Sampling Kafka and Data Sources # Kafka.

Behavioral changes

Learn about the change in certain functionality of Flink and Cloudera SQL Stream Builder that has resulted in a change in behavior from the previously released version of Cloudera Streaming Analytics - Kubernetes Operator.

Behavioral changes

Summary: [CSA-5372] - Use Apache Kerby KrbConfig for more robust EncType config, refactor keytab validation

Previous behavior:

The final Kerberos encryption type that is used by Cloudera SQL Stream Builder when generating a keytab was set according to the following precedence:

1. Values extracted from the keytab value given in the source.encytype.keytab Cloudera SQL Stream Builder configuration property.
2. Default values defined in the Cloudera SQL Stream Builder.

New behavior:

The final Kerberos encryption type that is used by Cloudera SQL Stream Builder when generating a keytab is set according to the following precedence:

1. Values defined in the krb5.conf file.
2. Values extracted from the keytab value given in the source.encytype.keytab Cloudera SQL Stream Builder configuration property.



Note: Note: This method is considered deprecated and will be removed in a future version of Cloudera Streaming Analytics.

3. Default values defined in the Cloudera SQL Stream Builder.

The following encryption types are the default in Cloudera SQL Stream Builder:

- aes256-cts-hmac-sha1-96
- aes128-cts-hmac-sha1-96
- des3-cbc-sha1
- arcfour-hmac
- des-hmac-sha1
- des-cbc-md5
- des-cbc-crc

Known issues and limitations

Learn about the known issues and limitations in this release of Cloudera Streaming Analytics - Kubernetes Operator.

Stuck session jobs in Cloudera Streaming Analytics - Kubernetes Operator

Session jobs stop running if the session cluster's Job Manager is restarted without High Availability configured. However, because of a Flink bug, such stopped jobs get stuck in RECONCILING/STABLE state and cannot be restarted or deleted.

In such cases, the following is seen when using the `kubectl get FlinkSessionJobs -n flink` command:

```
kubectl get FlinkSessionJobs -n flink
NAME                                JOB STATUS    LIFECYCLE STATE
ssb-ssbdefault-testjobname         RECONCILING   STABLE
```

1. Remove the FlinkDeployment resource related to Cloudera SQL Stream Builder:

```
kubectl -n [***NAMESPACE***] delete flinkdeployment [***DEPLOYMENT NAME***]
```

2. Open the Cloudera SQL Stream Builder configurations to edit:

```
kubectl -n [***NAMESPACE***] edit cm ssb-config
```

3. Add the highlighted line to the security-context.yaml entry:

```
allowPrivilegeEscalation: false
capabilities:
  drop:
  - ALL
runAsNonRoot: true
seccompProfile:
  type: RuntimeDefault
runAsUser: 9999
```

It can take a couple of minutes for the changes to take effect after modifying the ConfigMap.

4. Execute the SQL job from Cloudera SQL Stream Builder again.

FLINK-33536: S3 filesystem sink and CSV format throws error

When using the Flink Table API CSV streaming sink with the S3 filesystem, the operation fails with `IOException: Stream closed`.

Cloudera SQL Stream Builder

Flink

DataStream conversion limitations

- Converting between Tables and POJO DataStreams is currently not supported in Cloudera Streaming Analytics - Kubernetes Operator.
- Object arrays are not supported for Tuple conversion.
- The `java.time` class conversions for Tuple DataStreams are only supported by using explicit `TypeInfo`: `LegacyInstantTypeInfo`, `LocalTimeTypeInfo.getInfoFor(LocalDate/LocalDateTime/LocalTime.class)`.
- Only `java.sql.Timestamp` is supported for rowtime conversion, `java.time.LocalDateTime` is not supported.

Schema Registry catalog limitations

- Currently, the Schema Registry catalog / format only supports reading messages with the latest enabled schema for any given Kafka topic at the time when the SQL query was compiled.
- No time-column and watermark support for Registry tables.
- No CREATE TABLE support. Schemas have to be registered directly in the SchemaRegistry to be accessible through the catalog.
- The catalog is read-only. It does not support table deletions or modifications.
- By default, it is assumed that Kafka message values contain the schema id as a prefix, because this is the default behavior for the SchemaRegistry Kafka producer format.

To consume messages with schema written in the header, the following property must be set for the Registry client: `store.schema.version.id.in.header: true`.

Unsupported features

Learn what features are unsupported in this release of Cloudera Streaming Analytics - Kubernetes Operator.

Some Apache Flink and Cloudera SQL Stream Builder features exist in Cloudera Streaming Analytics - Kubernetes Operator, but are not supported by Cloudera. The following features are not ready for production deployment. Cloudera encourages you to explore these features in non-production environments and provide feedback on your experiences through the Cloudera Community Forums.

Flink

- Apache Flink batch (DataSet) API
- GPU Resource Plugin
- SQL Client
- The following features are not supported in SQL and Table API:
 - HBase Table Connector
 - Old Planner
 - Non-windowed (unbounded) joins, distinct

Cloudera SQL Stream Builder

- Materialized Views are not supported
- Virtual environments for Python are not supported
- Javascript UDFs are not supported

Deprecation notices in Cloudera Streaming Analytics - Kubernetes Operator

Certain features and functionalities have been removed or deprecated in Cloudera Streaming Analytics - Kubernetes Operator. You must review these items to understand whether you must modify your existing configuration. You can also learn about the features that will be removed or deprecated in the future release to plan for the required changes.

Terminology

Items in this section are designated as follows:

Deprecated

Technology that Cloudera is removing in a future Cloudera Streaming Analytics - Kubernetes Operator release. Marking an item as deprecated gives you time to plan for removal in a future Cloudera Streaming Analytics - Kubernetes Operator release.

Moving

Technology that Cloudera is moving from a future Cloudera Streaming Analytics - Kubernetes Operator release and is making available through an alternative Cloudera offering or subscription. Marking an item as moving gives you time to plan for removal in a future Cloudera Streaming Analytics - Kubernetes Operator release and plan for the alternative Cloudera offering or subscription for the technology.

Removed

Technology that Cloudera has removed from Cloudera Streaming Analytics - Kubernetes Operator and is no longer available or supported as of this release. Take note of technology marked as removed since it can potentially affect your upgrade plans.

Deprecation Notices for Cloudera SQL Stream Builder

Certain features and functionality are deprecated or removed in Cloudera SQL Stream Builder. You must review these changes along with the information about the features in Cloudera SQL Stream Builder that will be removed or deprecated in a future release.

Removed**v1 REST API**

The v1 REST API for Cloudera SQL Stream Builder is removed in Cloudera Streaming Analytics - Kubernetes Operator 1.4.0 and higher.

Cloudera recommends migrating to the v2 API.

For more information on the v2 API, see the [Cloudera SQL Stream Builder REST API reference](#).

Support for JavaScript UDFs

User-Defined Functions (UDFs) written in JavaScript are removed in Cloudera Streaming Analytics - Kubernetes Operator 1.4.0.

Cloudera recommends using Python UDFs and migrating JavaScript UDFs to Python. Support in Python UDFs for table transformations and the webhook connector is available for Cloudera SQL Stream Builder and Flink.

Component support

You can review the Cloudera Streaming Analytics - Kubernetes Operator components and their versions shipped in this release of the Cloudera Streaming Analytics - Kubernetes Operator.

Table 1: Cloudera Streaming Analytics - Kubernetes Operator component versions

Component	Version
Flink	1.20.1-csaop1.4.0-b55
Flink Operator	1.11-csaop1.4.0-b55
Cloudera SQL Stream Builder	1.20.1-csaop1.4.0-b55
OpenJDK	17.0.16

Supported Flink versions

Cloudera Streaming Analytics - Kubernetes Operator supports the following Flink versions:

Table 2: Supported Flink versions

Latest (default)	Other
1.20.1-csaop1.4.0-b55	None

The default version is the Cloudera-recommended current and latest supported Flink version. This version is used by default to deploy clusters if an explicit version is not provided in your `FlinkDeployment` resource.

Notice that the Flink versions are specific to Cloudera. Their version numbers consist of two parts: the first three digits specify the Apache Flink version, and the following specify the major and minor version of Cloudera Streaming Analytics - Kubernetes Operator. When deploying a cluster, you must use the Cloudera versions for Flink listed here. Specifying upstream versions is not supported.

Component support

Flink file system support

By default, Cloudera Streaming Analytics - Kubernetes Operator only supports local and locally mounted NFS/SAN file systems for Flink. Pluggable file systems (for example S3, HDFS, etc.) can be used by adding plugins to the Apache Flink Operator. For more information and a list of supported pluggable file systems, see *Using pluggable file systems*.

Cloudera SQL Stream Builder database support

You can use the following databases with Cloudera SQL Stream Builder:

- MySQL/MariaDB
- Oracle
- PostgreSQL

Cloudera SQL Stream Builder connector support

With Cloudera SQL Stream Builder, you will get the following connectors installed by default:

- Kafka
- JDBC
- CDC (MySQL, Oracle, Postgres, Db2, SqlServer)
- Amazon S3
- Azure Blob Storage
- Google Cloud Storage
- HDFS (when using the alternative image, see [Installation overview](#))
- Hive (when using the alternative image, see [Installation overview](#))
- Iceberg (when using the alternative image, see [Installation overview](#))
- Kudu (when using the alternative image, see [Installation overview](#))

You will also get the following formats installed with Cloudera SQL Stream Builder:

- JSON
- Avro
- ORC
- Parquet

Related Information

[Using pluggable file systems](#)

System requirements

Cloudera Streaming Analytics - Kubernetes Operator requires a Kubernetes cluster environment that meets the provided system requirements and prerequisites. You must ensure to meet these requirements to be able to install and use the Cloudera Streaming Analytics - Kubernetes Operator or its components.

- A Kubernetes 1.25 or later cluster:
 - OpenShift 4.12 or later
 - Any Cloud Native Computing Foundation (CNCF) certified Kubernetes distribution. For more information, see cncf.io.
- Administrative rights on the Kubernetes cluster.
- Access to kubectl or oc. These command line tools must be configured to connect to your running cluster.
- Access to helm.
- Log collection is enabled for the Kubernetes cluster.

Cloudera requires that the logs of Cloudera Streaming Analytics - Kubernetes Operator components are stored long term for diagnostic and supportability purposes. Review [Log collection](#).