

Cloudera Streaming Analytics - Kubernetes Operator 1.2.2 Release Notes

Date published: 2024-06-15

Date modified: 2024-02-28



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 1.2.2..... 4**
 - What's new..... 4**
 - Fixed issues..... 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 1.2.2

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 1.2.2

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) 1.9.0 and Apache Flink (Flink) 1.19.2. For more information about the supported versions, see the following upstream resources:

- [Flink Operator 1.9.0](#)
- [Flink 1.19.2](#)

Flink

Rebase to Flink 1.19.2

Apache Flink 1.19.2 is supported in Cloudera Streaming Analytics - Kubernetes Operator 1.2.2

For more information on what is included in the Apache Flink 1.19.2 version, see the [Apache Flink 1.19.2 Release Announcement](#) and [Release Notes](#).

Rebase to Flink Operator 1.9.0

Apache Flink Operator 1.9.0 is supported in Cloudera Streaming Analytics - Kubernetes Operator 1.2.2

For more information on what is included in the Apache Flink Operator 1.9.0 version, see the [Apache Flink 1.9.0 Release Announcement](#) and [Release Notes](#).

Cloudera SQL Stream Builder

General availability for Cloudera SQL Stream Builder

Cloudera SQL Stream Builder, previously in Technical Preview, is now generally available in Cloudera Streaming Analytics - Kubernetes Operator.

Cloudera SQL Stream Builder is a comprehensive interactive user interface for creating stateful stream processing jobs using SQL. For more information about Cloudera SQL Stream Builder and its features, see the [Getting started with Cloudera SQL Stream Builder](#) page.

LDAP authentication in Cloudera SQL Stream Builder

LDAP-based authentication is available for Cloudera SQL Stream Builder. For more information, refer to *LDAP authentication*.

Custom truststores

You can create custom truststores in the values.yaml file. For more information, refer to *Configuring custom truststores*.

Secure TLS connections to the Cloudera SQL Stream Builder UI and API

Users can connect to Cloudera SQL Stream Builder and API via a TLS encrypted channel. For more information, refer to *Routing with ingress*.

Using Python UDFs for table transformations and webhook connector

Due to the deprecation of Javascript for Cloudera Streaming Analytics - Kubernetes Operator (see [Deprecation notices](#)), support in Python UDFs for table transformations and the webhook connector have been added to Cloudera SQL Stream Builder and Flink.

Configurable requests and limits for all resources

Resource requests/limits can be configured in the values.yaml file. For more information, refer to *Resource requests and limits*.

Related Information

[LDAP authentication](#)

[Routing with ingress](#)

[Creating Python User-defined Functions](#)

[Resource requests and limits](#)

[Security configurations](#)

Fixed issues

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

CSA-5744 - Bump parquet libs to 1.15.1 in CSA Flink (backport FLINK-37609)

CSA-5441 - SSB can start an invalid job but can't stop it

CSA-5539 - [ssb][k8s] Postgresql pod crashes on Mirantis

CSA-5561 - [ssb] Fix server sent events indefinitely holding db connections

CSA-5569 - Delete flink kubernetes resource from SSB

CSA-5580 - Deletion of a running job is not handled properly.

CSA-5588 - Gracefully handle malformed UDFs

Improvements

CSA-4845 - Job Sampling in the CSA Operator deployments

CSA-5108 - Show default config values in job settings

CSA-5229 - Enhance Python editor in the UI

CSA-5318 - [ssb] Make webhook sink work with the sql-runner

CSA-5319 - [ssb] Make Kafka table transformation work with sql-runner

CSA-5468 - Upgrade CSA-OPERATOR images to latest UBI9

CSA-5491 - Backport Flink's OpenTelemetry Metrics Reporter

CSA-5501 - Expose pod level securityContext for postgres config

CSA-5503 - Ship CDP based "hadoop-common" in every FS connector to keep CDP - Operator compatibility

CSA-5518 - Specify resource requests/limits for all resources

CSA-5520 - Refactor schema detect UX for Kafka wizard tables

CSA-5529 - Rebase to flink-connector-kafka v3.3

CSA-5530 - [ssb][k8s] Make Kubernetes client timeout configurable

CSA-5540 - Add new flink-webhook-connector

CSA-5543 - Deprecate data transform in flink-connector-kafka

CSA-5585 - Minimize Flink docker image size with multistage build

CSA-5600 - Update warning text of job notifications in project home tab.

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 SSB:

```
kubectl -n [***NAMESPACE***] delete flinkdeployment [***DEPL
OYMENT NAME***]
```

2. Open the SSB 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

CSA-5622 [ssb][k8s] Project sync doesn't sync Kubernetes config for jobs

CSA-4858 - Kerberos encryption type detection does not always work correctly for Cloudera SQL Stream Builder

Cloudera SQL Stream Builder detects no supported encryption types even though there is a list of allowed encryption types in the `krb5.conf` file. This causes an error when generating keytabs from the principal and password pair.

1. Run `ktutil` on your cluster.
2. Change the configuration with the following commands:

```
addent -password -p [*** USERNAME ***] -k 1 -e aes256-cts  
wkt /tmp/new_keytab.keytab
```

3. Upload the new keytab on Streaming SQL Console.

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
- RAZ-enabled GCP environment

- 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
- Temporary views and tables are not supported
- Job names over 45 characters (including all unique tags, for example `ssb-[*** PROJECT NAME ***]-[*** JOB NAME ***]`) are not supported
- Virtual environments for Python 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.

Deprecated

v1 REST API

The v1 REST API for Cloudera SQL Stream Builder has been deprecated and will be removed in a future version of Cloudera Streaming Analytics - Kubernetes Operator.

Customers are advised to migrate to the v2 API, available for Cloudera SQL Stream Builder.

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

Support for JavaScript UDFs

Due to the deprecation of the Nashorn engine used in JDK 8 and 11, User-Defined Functions (UDFs) written in JavaScript has been deprecated in Cloudera Streaming Analytics - Kubernetes Operator. Cloudera recommends that customers start using [Python UDFs](#) for all new developments, and start migrating their JavaScript UDFs to Python to prepare for future upgrades.

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	1.19.2-csaop1.2.2-b6
Flink Operator	1.9-csaop1.2.2-b6
Cloudera SQL Stream Builder	1.19.2-csaop1.2.2-b6
OpenJDK	11.0.24

Supported Flink versions

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

Table 2: Supported Flink versions

Latest (default)	Other
1.19.2-csaop1.2.2-b6	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

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 1.23 or later cluster:
 - OpenShift 4.10 or later
 - RKE2 (Rancher Kubernetes Engine 2) 1.23 or later



Note: Cloudera Streaming Analytics - Kubernetes Operator complies with Cloud Native Computing Foundation (CNCF) standards and is compatible with CNCF-compliant Kubernetes distributions. For supporting your specific Kubernetes distribution, contact Cloudera.

- 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](#).