

Cloudera Streaming Analytics - Kubernetes Operator 1.1.1

## CSA Operator 1.1.1 Release Notes

Date published: 2024-06-15

Date modified: 2024-10-23

The Cloudera logo is displayed in a bold, orange, sans-serif font. The word "CLOUDERA" is written in all caps, with a stylized horizontal line through the middle of the letters 'E' and 'R'.

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

# Legal Notice

© Cloudera Inc. 2024. 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

<b>Release Notes.....</b>	<b>4</b>
What's new.....	4
Fixed issues.....	5
Known issues.....	5
Unsupported features.....	6
<b>Component support.....</b>	<b>6</b>
<b>System requirements.....</b>	<b>7</b>

## Release Notes

Learn about the new features, improvements, known and fixed issues, limitations, and unsupported features in this release of CSA Operator.

### What's new

Learn about the new features and notable changes in this release of CSA Operator.

#### CSA Operator 1.1.1

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

This release of the CSA Operator is based on the Apache Flink Kubernetes Operator (Flink Operator) 1.9.0 and Apache Flink (Flink) 1.19.1. For more information about the supported versions, see the following upstream resources:

- [Flink Operator 1.9.0](#)
- [Flink 1.19.1](#)

This version of CSA Operator also includes a Technical Preview of the SQL Stream Builder (SSB). SSB is a comprehensive interactive user interface for creating stateful stream processing jobs using SQL. For more information about SSB and its features, see the [What is SQL Stream Builder](#) page.

#### Flink

##### Rebase to Flink 1.19.1

Apache Flink 1.19.1 is supported in CSA Operator 1.1.0.

For more information on what is included in the Apache Flink 1.19.1 version, see the [Apache Flink 1.19.1 Release Announcement](#) and [Release Notes](#).

##### Rebase to Flink Operator 1.9.0

Apache Flink Operator 1.9.0 is supported in CSA Operator 1.1.0.

For more information on what is included in the Apache Flink Operator 1.9.0 version, see the [Apache Flink 1.19.1 Release Announcement](#) and [Release Notes](#).

#### SQL Stream Builder



**Note:** SQL Stream Builder is in Technical Preview in this release of CSA Operator.

##### Python UDFs in SQL Stream Builder

This feature allows customers to start using Python for creating User-Defined Functions (UDFs). 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, as Javascript UDFs will be removed in the future due to the deprecation of the Nashorn engine used in JDK 8 and 11.

For more information on using Python UDFs, see *Python UDFs*.

##### Global logging configuration for SQL Stream Builder jobs

A new global settings view has been enabled, which currently includes log4j configuration of Flink jobs started in SQL Stream Builder. Users with administrator rights in SQL Stream Builder can set a default logging configuration applied to all SSB jobs, which can be overridden at the job level.

For more information see *Adjusting logging configuration in Advanced Settings*.

### Customizable default Kafka TrustStore configuration in Streaming SQL Console

Customizing default Kafka TrustStore configurations was added to Streaming SQL Console. Kafka TrustStore can be configured during adding Kafka as a Data Source on the UI.

### Expose session cluster state and ability to stop it

SQL Stream Builder now displays information about existing session clusters and allows for those clusters to be terminated from the UI. For more information, refer to *TO-DO*.

### Related Information

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

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

## Fixed issues

Learn about the fixed issues in this release of CSA Operator.

**CSA-5065: Artifact Storage request thread does not timeout when storage is offline, hanging the UI**

**CSA-5291: [k8s][flink] The default Flink image version should point to the Cloudera repository**

**CSA-5306: SSB API does not validate catalog type**

**CSA-5356: Flink docker image permission issue**

**CSA-5424: Job Settings Logging tab bug**

### Improvements

**CSA-5145: [ssb] Only show connectors and formats which are available on the classpath**

**CSA-5278: [k8s][flink] Specify default USER in the Flink Dockerfile**

**CSA-5321 [ssb-ui] Hide Flink Dashboard from UI**

**CSA-5422: Use cloudbnative-pg image for postgres in the operator**

**CSA-5430: Update the Dockerfile in the flink-tutorials branch of the CSA Operator to use our own image**

## Known issues

Learn about the known issues in this release of CSA Operator.

### Stuck session jobs in Cloudera Streaming Analytics 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
```

Delete the jobs first, and then remove the finalizers from the object:

```
kubectl delete FlinkSessionJob [*** JOB NAME ***] -n flink
```

```
kubectl patch FlinkSessionJob [*** JOB NAME ***] -n flink --type
=merge -p '{"metadata":{"finalizers":[]}}'
```



**Important:** You need to do this for all stuck session jobs. You can only delete the `FlinkDeployment` and the namespace after all the stuck `FlinkSessionJobs` have been cleaned up.

### 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`.

## Unsupported features

Learn what features are unsupported in this release of CSA Operator.

The following upgrade scenarios are not supported in this version of CSA Operator:

- Upgrading from the Technical Preview version to the Generally Available version
- Upgrading from the Apache Flink Kubernetes Operator to CSA Operator

This version of the CSA Operator also includes the Technical Preview of SQL Stream Builder (SSB). The following SSB features are not included in this version of CSA Operator:

- Materialized Views
- Notifications
- Sampling
- Webhook sinks
- Kafka data transformations

## Component support

You can review the CSA Operator components and their versions shipped in this release of the CSA Operator.

**Table 1: CSA Operator component versions**

Component	Version
Flink	1.19.1-csaop1.1.1-b13
Flink Operator	1.9-csaop1.1.1-b13
SQL Stream Builder	1.19.1-csaop1.1.1-b13
OpenJDK	11.0.24

### Supported Flink versions

CSA Operator supports the following Flink versions:

**Table 2: Supported Flink versions**

Latest (default)	Other
1.19.1-csaop1.1.1-b13	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 CSA 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, CSA 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*.

#### SSB database support

You can use the following databases with SQL Stream Builder (SSB):

- MySQL/MariaDB
- Oracle
- PostgreSQL

#### SSB connector support

With SSB, 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 SSB:

- JSON
- Avro
- ORC
- Parquet

## System requirements

CSA 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 CSA Operator or its components.

- A Kubernetes 1.23 or later cluster with OpenShift 4.10 or later.
- 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 CSA Operator components are stored long term for diagnostic and supportability purposes. Review [Log collection](#).