

Cloudera Streaming Analytics 1.8.0

Release Notes

Date published: 2019-12-17

Date modified: 2022-09-28

CLOUDERA

<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

What's new in Cloudera Streaming Analytics.....	4
Streaming SQL Console improvements.....	5
Fixed issues.....	7
Known issues and limitations.....	7
Behavioral changes.....	10
Unsupported features.....	10
Component support.....	11

What's new in Cloudera Streaming Analytics

Cloudera Streaming Analytics 1.8.0 covers new features beside the core streaming functionality of Apache Flink and SQL Stream Builder.

CSA 1.8.0.3

Hotfix for CSA 1.8.0

On February 2, 2023, CSA 1.8.0.3 was released as a hotfix to resolve some issues. If you have previously upgraded to CSA 1.8.0.0, 1.8.0.1 or 1.8.0.2 upgrade to CSA 1.8.0.3 to access the latest version of CSA. You can find the new download links in the [Download location of CSA](#).

For more information about the issues fixed in CSA 1.8.0.3, see the [Fixed issues](#) section.

CSA 1.8.0.2

Hotfix for CSA 1.8.0

On December 19, 2022, CSA 1.8.0.2 was released as a hotfix to resolve some issues. If you have previously upgraded to CSA 1.8.0.0 or 1.8.0.1, upgrade to CSA 1.8.0.2 to access the latest version of CSA. You can find the new download links in the [Download location of CSA](#).

For more information about the issues fixed in CSA 1.8.0.2, see the [Fixed issues](#) section.

CSA 1.8.0.1

Hotfix for CSA 1.8.0

On October 19, 2022, CSA 1.8.0.1 was released as a hotfix to resolve some issues. If you have previously upgraded to CSA 1.8.0.0, upgrade to CSA 1.8.0.1 to access the latest version of CSA. You can find the new download links in [Download location of CSA](#).

For more information about the issues fixed in CSA 1.8.0.1, see the [Fixed issues](#) section.

CSA 1.8.0.0

Apache Flink upgrade

Apache Flink 1.15.1 is supported in Cloudera Streaming Analytics 1.8.0.

For more information on what is included in the Apache Flink 1.15 version, see the [Apache Flink 1.15.1 Release Post](#) and the [Apache Flink 1.15 Release Notes](#)

Job notifications

You can enable notifications for SQL jobs to send information about job status changes indicated under the **Events** tab.

For more information, see the [Using SQL job notification](#) documentation.

Diagnostic data collection

Data collection for diagnostic information is enabled for SQL Stream Builder to make troubleshooting and support processes easier.

For more information, see the [Collecting diagnostic data](#) documentation.

Migrating SQL jobs

You can import and export SQL jobs directly using the Streaming SQL Console for easier job migration from testing environments to production, or after upgrading CSA.

For more information, see the [Migrating SQL jobs](#) documentation.

Materialized View pagination

You can set a limit and offset for the results of Materialized View queries to filter the query results more easier when accessing them through the Query API.

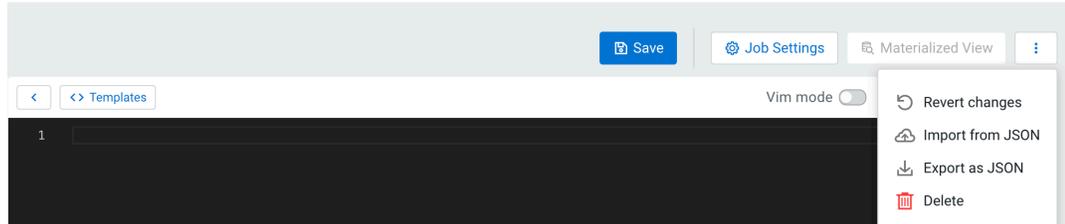
For more information, see the [Materialized View Pagination](#) documentation.

Streaming SQL Console improvements

Learn more about the improvements and changes for the User Interface (UI) of SQL Stream Builder in the Cloudera Streaming Analytics 1.8.0 release.

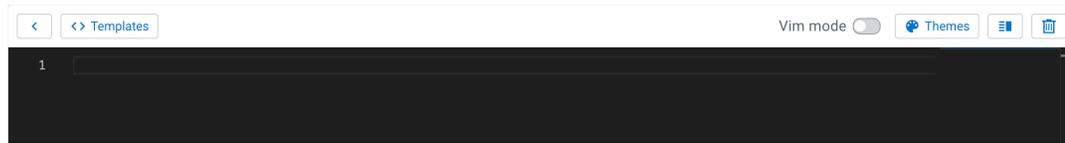
Job management buttons

The Revert and Delete button have been moved to a new location on the **Console** page:



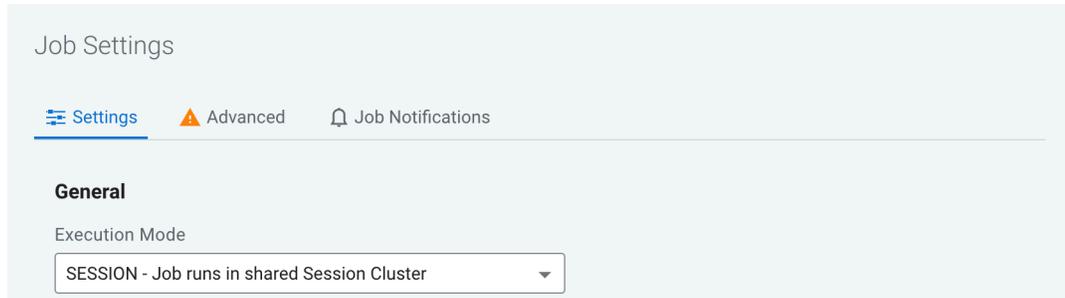
Reorganized upper panel of SQL Editor

The elements on the upper panel of **SQL Editor** have been reorganized:



Added selector for job execution mode

The execution mode of jobs can be selected in **Job Settings** window:



Showing separate results when executing multiple statements

When executing more than one SQL statements, the results can be viewed in a vertical layout:

The screenshot shows a SQL query editor with the following code:

```

1 show tables;
2 desc datagen_table_1664297999;
3 select * from datagen_table_1664297999;
    
```

Below the editor, the **Results** tab is active, displaying a table with the following data:

col_str	col_int	col_ts
880d4a2604645ce680cec21eb4958...	123617109	2022-09-28T10:06:02.798
abc2bb8680a1fddd55a6f39c6db62...	165237774	2022-09-28T10:06:03.802

Autocompleting column names

The column name is automatically completed when adding a dot after a table name:

The screenshot shows the SQL editor with the following code:

```

1 SELECT datagen_table_1664297999. from datagen_table_1664297999
    
```

An autocomplete dropdown menu is visible, listing the columns: `col_int`, `col_str`, and `col_ts`.

Selecting all or individual sample rows with short keys

You can use space to select individual sample rows or CTRL/Command + A to select all rows under **Results** tab.

The screenshot shows the **Results** tab with the following data, where all rows are selected (indicated by blue checkmarks):

col_str	col_int	col_ts
86be6a41f489667f42887e8fd7817758631...	1763709221	2022-09-28T10:17:28.260
b717de9bf4b6ff074262f04bc239c3516ea4f...	1495541120	2022-09-28T10:17:24.257
fb5aa9713743fccfca5e4c0930d3de3ba7e8...	991595962	2022-09-28T10:17:23.256
737b1101965e983de87fe1cda5f40e87b99f...	704934579	2022-09-28T10:17:15.242
de77361457a76f776603a2523bf936b0d0a9...	601713056	2022-09-28T10:17:20.247

Fixed issues

Review the list of Flink and SQL Stream Builder issues that are resolved in Cloudera Streaming Analytics 1.8.0.

February 2, 2023

CSA-4390: Missing flink-shaded-zookeeper-3 artifacts

The missing flink-shaded-zookeeper-3 artifacts have been added to the public artifact repository of CSA 1.8.0.

December 19, 2022

CSA-4358: Compatibility issue for CDP Private Cloud Base 7.1.7 and Flink 1.15

The compatibility issue regarding missing Flink filesystem connector dependencies between CDP Private Cloud Base 7.1.7, 7.1.7 Service Pack (SP) 1 and Flink 1.15 has been fixed.

October 19, 2022

SQL Stream Builder

CSA-4047: HSTS headers are not added to the response

The issue about Flink History Server not adding the HSTS headers has been fixed.

CSA-4021: SSB Load Balancer fails to start on a kerberized cluster

The issue about Load Balancer failing when using a kerberized cluster is resolved. Support for custom SSL key and certificate to Load Balancer has been added.

Flink

FLINK-16635: Dependency for okio and okhttp from flink-metrics-influxdb

The pinned dependency for okio and okhttp from flink-metrics-influxdb has been removed.

September 29, 2022

SQL Stream Builder

CSA-3529: Filtering ARRAY typed columns do not work for Materialized Views

The issue about ARRAY typed columns are treated as STRING when setting filters for a Materialized View query has been fixed.

CSA-3507: IN and NOT IN operators cannot be used for Materialized Views

The issue about using IN and NOT IN operators when filtering Materialized Views has been fixed.

Known issues and limitations

Learn about the known issues in Flink and SQL Stream Builder, the impact or changes to the functionality, and the workaround in Cloudera Streaming Analytics 1.8.0.

SQL Stream Builder

FLINK-18027: ROW value constructor cannot deal with complex expressions

When querying data from a table or a view with a ROW() function an exception is thrown due to a Calcite parsing issue. For example, the following query will return an error:

```
CREATE VIEW example AS SELECT col1, ROW(col2) FROM table;
```

```
SELECT * FROM example;
```

Add a second SELECT layer to the SQL query as shown in the following example:

```
CREATE VIEW example AS SELECT col1, ROW(col2) FROM (SELECT col1,
col2 FROM table);
SELECT * FROM example;
```

Auto discovery is not supported for Apache Knox

You need to manually configure Knox with SQL Stream Builder to enable Knox authentication.

Complete the configuration based on the CDP Private Cloud Base version you use. For more information, see the [Enabling Knox authentication for SSB](#) documentation.

Streaming SQL Console cannot be accessed through Knox when High Availability is enabled

When SQL Stream Builder SSB is deployed in High Availability with Load Balancer, the Streaming SQL Console cannot be accessed directly using Apache Knox.

For accessing Streaming SQL Console, use the secured Load Balancer deployment or authenticate using SPNEGO.

CSA-5006 - SSB service fails when using Active Directory (AD) Kerberos authentication

If you use AD Kerberos for authentication and the Load Balancer URL is not provided, it can cause the SQL Stream Builder (SSB) service to fail. The issue is caused by the keytab generation. When the keytab is generated by Cloudera Manager it requires the principals from the AD for the Load Balancer host, and without no host specified for the Load Balancer, the SSB service cannot be started by Cloudera Manager. This issue also persists when the Load Balancer role is not deployed or used with SSB.

Fill out the Load Balancer URL parameter in Cloudera Manager regardless of using Load Balancer with SSB. For more information, see the [Enabling High Availability for SSB](#) documentation.

CSA-4425: Password in Kafka Data Source can be revealed after save

The show password icon can be used after saving the password for the authentication method when creating a Kafka Data Source.

None

CSA-4333: Use Kafka Timestamps switch reflects invalid value

After creating a Kafka Virtual Table and disabling the Use Kafka Timestamps configuration, the table is created successfully according to the setting, but when viewing the DDL of the table, it shows the configuration as enabled.

None

CSA-4030: Webhook sending fails when webhook template is empty string

When creating a webhook table with a custom template, the webhook template will be saved as an empty string, which results in webhook sending failure.

None

CSA-3754: The display name of the loadbalancer.url property should be "Load Balancer Host"

The loadbalancer.url property is duplicated in Cloudera Manager on the SQL Stream Builder configuration page.

None

CSA-2016: Deleting table from other teams

There is a limitation when using the Streaming SQL Console for deleting tables. It is not possible to delete a table that belongs to another team using the Delete button on the User Interface.

Use DROP TABLE statement from the SQL window.

CSA-1454: Timezone settings can cause unexpected behavior in Kafka tables

You must consider the timezone settings of your environment when using timestamps in a Kafka table as it can affect the results of your query. When the timestamp in a query is identified with `from_unixtime`, it returns the results based on the timezone of the system. If the timezone is not set in UTC+0, the timestamp of the query results will shift in time and will not be correct.

Change your local timezone settings to UTC+0.

CSA-1231: Big numbers are incorrectly represented on the Streaming SQL Console UI

The issue impacts the following scenarios in Streaming SQL Console:

- When having integers bigger than 253-1 among your values, the Input transformations and User Defined Functions are considered unsafe and produce incorrect results as these numbers will lose precision during parsing.
- When having integers bigger than 253-1 among your values, sampling to the Streaming SQL Console UI produces incorrect results as these numbers will lose precision during parsing.

None

Flink

FLINK-18027: ROW value constructor cannot deal with complex expressions

When querying data from a table or a view with a `ROW()` function an exception is thrown due to a Calcite parsing issue. For example, the following query will return an error:

```
CREATE VIEW example AS SELECT col1, ROW(col2) FROM table;
SELECT * FROM example;
```

Add a second `SELECT` layer to the SQL query as shown in the following example:

```
CREATE VIEW example AS SELECT col1, ROW(col2) FROM (SELECT col1,
col2 FROM table);
SELECT * FROM example;
```

In Cloudera Streaming Analytics, the following SQL API features are in preview:

- Match recognize
- Top-N
- Stream-Table join (without rowtime input)

DataStream conversion limitations

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

Kudu catalog limitations

- `CREATE TABLE`
 - Primary keys can only be set by the `kudu.primary-key-columns` property. Using the `PRIMARY KEY` constraint is not yet possible.
 - Range partitioning is not supported.
- When getting a table through the catalog, `NOT NULL` and `PRIMARY KEY` constraints are ignored. All columns are described as being nullable, and not being primary keys.
- Kudu tables cannot be altered through the catalog other than simply renaming them.

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 behaviour 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`.

Behavioral changes

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

SQL Stream Builder

Summary:

Keytabs can only be uploaded for new users

Previous behavior:

A new keytab could be uploaded for every user even though they already uploaded and unlocked their keytab on Streaming SQL Console.

New behavior:

New users can upload their keytab to Streaming SQL Console. Already registered users need to lock the keytab and update their credentials to use a newer keytab version on Streaming SQL Console.

Flink

Summary:

Additional step required when upgrading to latest version of CSA

Previous behavior:

Creating Job Result Store was not needed when upgrading to latest version of CSA.

New behavior:

When upgrading to latest version of CSA, you need to create the Job Result Store for Flink.

Unsupported features

The following features are not supported in Cloudera Streaming Analytics 1.8.0.

SQL Stream Builder

- Direct SQL Stream Builder upgrade from 1.3.0



Important: This does not impact Flink, you can directly upgrade Flink as described in the documentation.

For more information, see the [Upgrading SQL Stream Builder](#) in the 1.3.0 documentation.

Flink

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

Component support

Learn more about which Apache Flink component version is supported in the Cloudera Streaming Analytics (CSA) releases.

CSA version	Component version
CSA 1.8.0	Apache Flink 1.15.1
CSA 1.7.0	Apache Flink 1.14
CSA 1.6.2	
CSA 1.6.1	
CSA 1.6.0	
CSA 1.5.3	Apache Flink 1.13
CSA 1.5.1	
CSA 1.5.0	
CSA 1.4.1	Apache Flink 1.12
CSA 1.4.0	
CSA 1.3.0	
CSA 1.2.0	Apache Flink 1.10
CSA 1.1.0	Apache Flink 1.9.1

Related Information

[CSA 1.7.0 Release Notes](#)

[CSA 1.6.2 Release Notes](#)

[CSA 1.6.1 Release Notes](#)

[CSA 1.6.0 Release Notes](#)

[CSA 1.5.3 Release Notes](#)

[CSA 1.5.1 Release Notes](#)

[CSA 1.5.0 Release Notes](#)

[CSA 1.4.1 Release Notes](#)

[CSA 1.4.0 Release Notes](#)

[CSA 1.3.0 Release Notes](#)

[CSA 1.2.0 Release Notes](#)

[CSA 1.1.0 Release Notes](#)