

Cloudera Streaming Analytics 1.6.0

Release Notes

Date published: 2019-12-17

Date modified: 2021-12-15

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

CVE-2021-45105 & CVE-2021-44832 remediation for CSA.....	4
What's new in Cloudera Streaming Analytics.....	4
Fixed issues.....	5
Component support.....	5
Known issues and limitations.....	6
Unsupported features.....	8

CVE-2021-45105 & CVE-2021-44832 remediation for CSA

Learn more about the CVE-2021-45105 and CVE-2021-44832 remediation for Flink and SQL Stream Builder in Cloudera Streaming Analytics (CSA).

Cloudera released maintenance versions for CSA on CDP Private Cloud Base to address CVE-2021-45105 and CVE-2021-44832 identified as critical vulnerability issues for Log4j2.

Cloudera encourages users to upgrade to the following CSA versions to avoid any possibility of exploitation:

- [CVE-2021-45105](#): CSA 1.5.3 or higher version
- [CVE-2021-44832](#): CSA 1.6.1

For more information about the impacts of CVE-2021-45105, see the [TSB 2021-547: Critical vulnerability in log4j2 CVE-2021-45105](#) Knowledge Base article.

Related Information

[CSA 1.5.3 Installation guide](#)

[CSA 1.6.1 Installation guide](#)

What's new in Cloudera Streaming Analytics

Cloudera Streaming Analytics 1.6.0 covers new features beside the core streaming functionality of Apache Flink and SQL Stream Builder.

Apache Flink upgrade

Apache Flink 1.14 is supported in Cloudera Streaming Analytics 1.6.0.

For more information on what is included in the Apache Flink 1.14 version, see the [Apache Flink 1.14 Release Post](#) and the [Apache Flink 1.14 Release Notes](#).

Db2 CDC connector support

Db2 CDC connector is added for the set of supported connectors. This enables you to use data from the IBM Db2 databases with the Change Data Capture connector in SSB.



Note: The Db2 connector is only supported for Linux operating system.

For more information, see the [CDC connectors documentation](#).

Custom connectors and data formats

You are able to add new connectors and data formats, and also modify the default and predefined connectors and data formats using Streaming SQL Console to further customize your Streaming SQL jobs.

For more information, see the *Supported connectors* and *Managing connectors* in the [Connectors](#) sections.

SQL job migration tool

With the SQL job migration tool you can migrate your SQL jobs with configurations from one cluster to another using REST API or Command Line Interface (CLI).

For more information, see the [Using the SQL job migration tool](#) section.

REST API extended support

The SQL Stream Builder REST API is extended with more job, session and query management operations for Flink and SSB. There is also a direct link for the REST API Explorer from Cloudera Manager and from Streaming SQL Console as well.

For the list of newly added operations and more information about the SSB REST API, see the [Using the SSB REST API](#) section.

Flink Dashboard access

Flink Dashboard is added to the Streaming SQL Console main menu for easier job monitoring access. The SSB Team authorization feature is also extended to Flink Dashboard which means that only those jobs can be monitored in Flink Dashboard that belong to the team a user is assigned to.

Fixed issues

Review the list of Flink and SQL Stream Builder issues that are resolved in Cloudera Streaming Analytics 1.6.0.

CSA-2389: Describing Kafka table fails when array is nested in schema

The issue about describing a Kafka table when an arrays is nested within the schema is fixed.

CSA-2321: SASL/PLAINTEXT is not available

The issue regarding the SASL/PLAINTEXT protocol not being available when using Kafka as a data provider for SQL Stream Builder is fixed.

CSA-2301: Transformation error when Kafka has null values

The issue regarding the transformation error when there are null values in Kafka messages is fixed.

CSA-1985: DROP TABLE limitation when using Webhook table

The issue regarding the execution of DROP TABLE statement when using Webhook table is fixed.

Component support

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

CSA version	Component version
CSA 1.6.0	Apache Flink 1.14
CSA 1.5.1	Apache Flink 1.13
CSA 1.5.0	Apache Flink 1.13
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 1.5.3 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](#)

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 1.6.0.

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;
```

Cannot access API Explorer

The **API Explorer** page of SSB REST API cannot be accessed when using Apache Knox as authentication method. This issue is not present when using SPNEGO authentication.

None

Uploading connector files fail

When trying to upload a new connector JAR with a file size more than 1 MB, the upload process fails with an error.

Set the `server.tomcat.max-swallow-size` in Cloudera Manager using the following steps:

1. Open your cluster in Cloudera Manager.
2. Select SQL Stream Builder from the list of services.
3. Select Configuration.
4. Search for Streaming SQL Engine Advanced Configuration Snippet (Safety Valve) for `ssb-conf/application.properties` in the search bar.
5. Add `server.tomcat.max-swallow-size=2000MB` to the **Safety Valve**.
6. Click Save.
7. Restart the SQL Stream Builder service.

CSA-2559: Materialized View settings can be overwritten while running job

Materialized View settings are overwritten when submitting a new job with the same name.

None

CSA-2551: Dynamic filters are not working with greater value than a character

The dynamic filtering feature cannot be used for the Materialized View when providing a parameter with a value that is greater than the value of a character type.

CSA-2547: Vulnerability issue for user impersonation

With SPENGO authentication, adding the `doAs=other_user` parameters results in that users can be impersonated as the call is proxied to the Streaming SQL Engine as `ssb` principal.

CSA-2538: Error when saving Materialized View configuration

Due to data type mismatch for `retention_interval_ms` in the console and admin databases, configurations using MV retention times (greater than 2147483 seconds) cannot be saved for Materialized Views.

CSA-2529: Cannot set consumer groups for Kafka tables

Queries fail when adding consumer groups for Kafka table settings.

CSA-2528: Improvement for Materialized View table names

The automatically created names of Materialized View tables are not expressive enough to easily work with.

Db2 CDC connector is not available from Connectors and Templates

The Db2 Change Data Capture (CDC) is not yet available on the Streaming SQL Console under Templates and Connectors. This does not limit the use of the Db2 connector.

You can use the following the Db2 CDC example as a reference to create a table:

```
CREATE TABLE db2_cdc_source (  
  'column_name' INT,  
  'column_name' STRING  
) WITH (  
  'connector' = 'db2-cdc',  
  'hostname' = '...',  
  'port' = '...',  
  'username' = '...',  
  'password' = '...',  
  'database-name' = '...',  
  'schema-name' = '...',  
  'table-name' = '...'  
)
```

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-1673: SSB operations are not showing in Atlas

Due to a communication issue SQL Stream Builder (SSB) operations are not showing in Atlas.

None

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.

Unsupported features

The following features are not supported in Cloudera Streaming Analytics 1.6.0.

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