

Cloudera Streaming Analytics 1.6.2

Release Notes

Date published: 2019-12-17

Date modified: 2022-02-28

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
Fixed issues.....	4
Component support.....	4
Known issues and limitations.....	5
Unsupported features.....	7

What's new in Cloudera Streaming Analytics

Cloudera Streaming Analytics 1.6.2 covers new features beside the core streaming functionality of Apache Flink and SQL Stream Builder.

Configuring deserialization failures for Kafka tables

Support for configuring error handling of deserialization is added. When using the Kafka connector with any data type, you can choose from the following options how to handle schema mismatch error:

- Throw an exception
- Ignore the message
- Ignore the message and log the error
- Ignore the message in the context of the current stream, but store it in a dead-letter queue topic

Fixed issues

Review the list of Flink and SQL Stream Builder issues that are resolved in Cloudera Streaming Analytics 1.6.2.

CSA-2708: SQL Stream Builder (SSB) service fail to start on Ubuntu20

The issue of SSB not start when using them on Ubuntu20 has been fixed.

CSA-2659: Catalogs are removed after SSB restart

The issue of registered catalogs being removed after the SQL Stream Builder (SSB) service is restarted has been fixed.

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

The issue of Materialized View settings are overwritten when submitting a new job with the same name has been fixed.

CSA-1673: SSB operations are not showing in Atlas

The issue of SQL Stream Builder (SSB) operations not showing in Atlas has been fixed. Atlas can be used to monitor SSB jobs.

Component support

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

CSA version	Component version
CSA 1.6.2	Apache Flink 1.14
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 version	Component version
CSA 1.2.0	Apache Flink 1.10
CSA 1.1.0	Apache Flink 1.9.1

Related Information

[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](#)

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 1.6.2.

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 size file 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-3537: Catalogs are deleted after Streaming SQL Engine restart

When the Streaming SQL Engine instance is restarted on the cluster, the already registered catalogs are removed.

None

CSA-3464: Overwriting new changes for SQL Job with edit

A previous version of a job setting is applied when the SQL query is changed for an existing job after reloading it to the **Compose** page. The issue can be experienced when clicking on the Edit Job button on the **SQL Jobs** page.

None

CSA-3308: Changing Primary Key for Materialized View without recreating table causes job failure

The SQL job fails when directly changing the primary key of Materialized View and restarting the SQL job due to schema mismatch. The SQL job silently keeps retrying and overloads the Postgres database when the primary key of a Materialized View got changed without deleting existing data.

None

CSA-2729: DLQ topic is filled with unexpected results

Some messages that are left in the buffer can be sent to the Dead Letter Queue (DLQ) topic after handling a deserialization failure for Kafka tables.

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;
```

FLINK-27441: Scrollbar is missing for particular UI elements on Flink Dashboard

The angular version bump introduced a bug that causes the scrollbar on the Flink Dashboard invisible.

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 1.6.2.

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