Cloudera Streaming Analytics 1.6.1

Release Notes

Date published: 2019-12-17
Date modified: 2022-01-19

CLOUDERA

https://docs.cloudera.com/

Legal Notice

Contents

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

As the Cloudera Streaming Analytics 1.6.1 is a maintenance release, there are no new features available. For more information about the issues fixed in the maintenance release, see the *Release Notes - Fixed issues* section.

Fixed issues

Review the list of Flink and SQL Stream Builder issues that are resolved in Cloudera Streaming Analytics 1.6.1.

CSA-2566: Adding template support for Db2 CDC connector

Template is available for the Change Data Capture (CDC) Db2 connector in Streaming SQL Console.

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

CSA-2564: Log4j version update

The Log4j version is updated to 2.17.1 to include the fix for CVE-2021-44832.

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

The issue of the dynamic filters of Materialized View is fixed. The endpoint works properly even with greater values than a character type.

CSA-2547: Vulnerability issue for user impersonation

The vulnerability issue of using the doAs=other_user parameters is fixed. Users cannot be impersonated when using SPNEGO authentication.

CSA-2538: Error when saving Materialized View configuration

The issue of saving configuration of Materialized View is fixed. Materialized View retention times are now essentially unlimited (with a maximum of seconds represented by a 64-bit integer).

CSA-2529: Cannot set consumer groups for Kafka tables

The issue of setting Consumer groups for Kafka tables are fixed. Queries can be executed even when Consumer groups are specified.

CSA-2528: Improvement for Materialized View table names

The naming of Materialized View tables has been improved to allow for user configurable, per-organization prefixes for better organization.

Component support

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

CSA version	Component version
CSA 1.6.1	Apache Flink 1.14
CSA 1.6.0	Apache Flink 1.14
CSA 1.5.3	Apache Flink 1.13
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 1.6.0 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 1.6.1.

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-2708: SSB service fail to start on Ubuntu20

When using SQL Stream Builder (SSB) on Ubuntu20, the SSB service fails to start due to missing Python references.

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-2659: Catalogs are removed after SSB restart

Registered catalogs are removed after the SQL Stream Builder (SSB) service is restarted.

None

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-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-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-1985: DROP TABLE limitation when using Webhook table

DROP TABLE cannot be executed against Webhook type tables. The following error message is displayed when trying to delete a Webhook table using the SQL window: Table with identifier 'xyz' does not exist.

Use the Delete button on the Streaming SQL Console.

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-1232: 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 PRIM ARY 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 1.6.1.

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