

Using Materialized Views

Date published: 2019-12-17

Date modified: 2021-07-20



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

Using Materialized Views.....	4
Creating Materialized Views.....	4

Using Materialized Views

SQL Stream Builder has the capability to materialize results from a Streaming SQL query to a persistent view of the data that can be read through REST. Business Intelligence tools and applications can use the Materialized View REST endpoint to query streams of data without deploying database systems.

Materialized Views are in synchronization with the mutating stream - they are updated by a primary key as data flows through the system. The data is updated by a given key, and it represents the latest view of the data by key.

For example: vehicleID Z latest latitude and longitude is X and Y. As the vehicle moves, the latitude and longitude for the vehicleID are updated. The primary key is defined at creation time and is immutable.

Materialized Views can be created as mutating snapshots of the queried data result that is updated by a given key. The data is always the latest representation of itself by key (analogous to a primary key in most RDBMS systems).

You can query the Materialized Views using a GET request over REST, which returns a JSON response as "Content-Type: application/json". The queries are not defined at query time. Rather, they are curated, saved, and granted access through the Cloudera platform. You can configure a REST endpoint to query the Materialized View. Multiple query conditions can be created to allow various ways to query the same data. This is sometimes referred to as a 'pull query'.

Null Keys

When a key is removed from the incoming messages of a source, SSB continues to consume them. However, it marks the missing key as NULL at the sink. Similarly, when a key is removed from the source schema, but not from the incoming messages of the source, SSB ignores the key on the incoming stream.

Creating Materialized Views

After executing a SQL Stream job, you can set up the Materialized Views to have a snapshot of your queried data. You can use the URL Pattern from the Materialized View to visualize the generated data.

Before you begin

- Make sure that PostgreSQL is installed and configured to SQL Stream Builder (SSB) to create Materialized Views.
- You need to have a running SQL job on which you apply the Materialized Views configuration.

Procedure

1. Go to your cluster in Cloudera Manager.
2. Click SQL Stream Builder from the list of services.
3. Click SQLStreamBuilder Console.
The Streaming SQL Console opens in a new window.
4. Run a SQL Stream job, and wait until data is shown on the results tab.

For more information on running a SQL job, see the Running a SQL Stream job documentation.

5. Select the Materialized View sub-tab on the Compose tab.

The screenshot shows the 'Materialized View' configuration interface. At the top, there are two tabs: 'SQL' and 'Materialized View', with 'Materialized View' being the active tab. Below the tabs is a 'Configuration' section. It contains several settings: 'Primary Key' with a dropdown menu showing 'No keys available'; 'Retention (Seconds)' with a text input field containing '300'; 'Recreate on Job Start' with a checked checkbox; 'Ignore NULLs' with an unchecked checkbox; and 'API Key' with a dropdown menu showing 'Select a key'. A red error message 'Please select or create an API Key' is displayed below the API Key dropdown. To the right of the configuration settings is a blue informational box that reads: 'Enter a valid query in the "Streaming Query" tab to begin building your materialized view.' Below the configuration section is a 'Materialized View Queries' section. It contains a large blue box with a list of instructions: 'Enter a valid SQL query', 'Select or Create an API Key', and 'Click Apply Configuration to commit configuration.' To the right of this box is a green 'Add Query' button.

6. Select a Primary Key.

If this list is empty, then no SQL is specified on the SQL sub-tab, or that SQL is invalid. Select a key as a primary key for the Materialized View. All data will be updated by this key.

7. Select a Retention Period.

Data not being mutated during this period is removed from the view.

8. Enable or disable Recreate on Job Start.

If enabled, the Materialized View is deleted when a job is started or restarted.

9. Enable or disable Ignore NULLS.

If enabled, NULL values will NOT update values that are non null - they are ignored.

10. Select an API Key.

In case there are no API Keys, click Add API key, or click Materialized Views from the main menu. The add API key window appears. Provide a name for the API key, and click Save Changes.

API Key

Key Value

640bef50-ba81-44c7-bce9-140faa2ad7d6

Key Name

API Key Name

Please provide an API Key name.

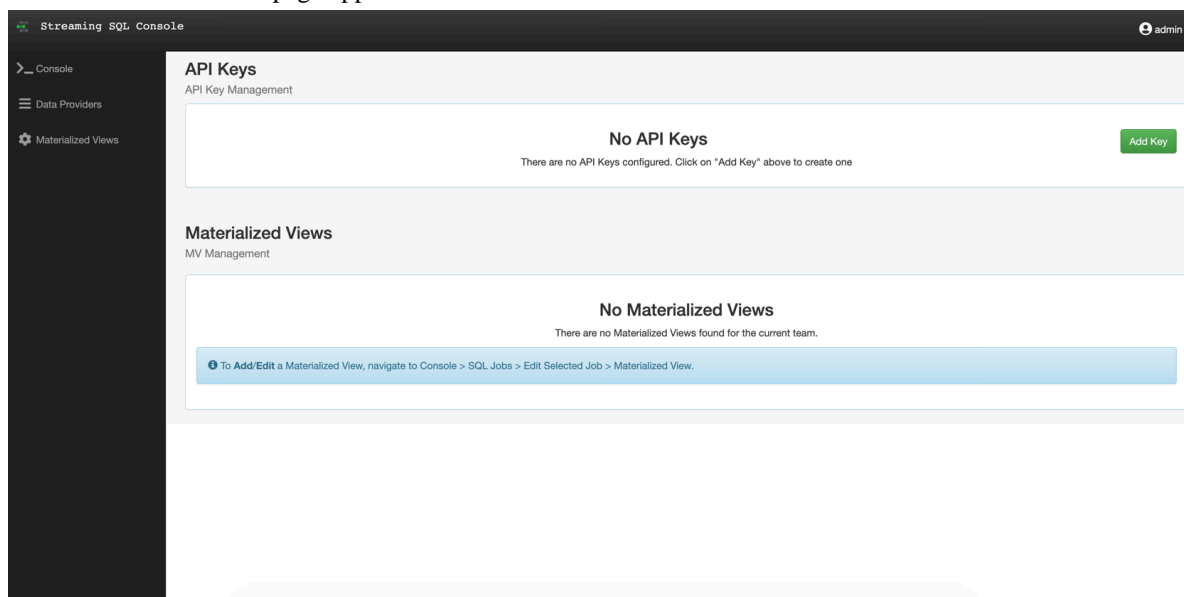
Close

Save Changes

To check your created API keys:

- a. Click Materialized Views from the main menu.

The Materialized Views page appears.



- b. Click [show] under Key to see the API Key.



Note: You can delete the created API key by clicking the bin icon.

11. Click Apply Configuration.

12. Click Add Query to create the Materialized View query.

The Materialized View Query Configuration window appears.

Materialized View Query Configuration

URL Pattern

/api/v1/query/<Job ID>/

This is the API Endpoint pattern. Denote parameters inside curly braces, i.e. "{param}"

Description (optional)

Query Builder

name



Add Column

Select All

Unselect All

No Data Selected

Filters

Add Filters

No Filters Configured

Close

Save Changes

13. Provide a name to the URL Pattern.

14. Provide a description of the Materialized View Query, if needed.

15. Customize the Materialized View in the Query Builder.


Query Builder

name

Add Column

Select All

Unselect All

Name	Alias	Type	Actions
name	name	VARCHAR	

Filters

AND

OR

+ Add rule

+ Add group

name

equal

test

✕ Delete

Close

Save Changes

- Select the columns of the SQL job you want to use in the Materialized View Query.
- Click Add filters to apply computations and further enrichment of your data.
- Click Add rule to save the filter.

16. Click Save Changes.

The Pattern URL appears under Materialized View Queries header.

Materialized View Queries

URL Pattern	Description	Add Query
/api/v1/query/5199/docstest?key=640bef50-ba81-44c7-bce9-140faa2ad7d6		 

Results

You can click the created REST endpoint to review the data, or copy it and visualize the queried data in a Business Intelligence tool, notebook, code and so on.

Related Information

[Configuring databases for SSB](#)

[Running a SQL Stream job](#)