Enable SSO for JDBC/ODBC Connections to Virtual Warehouses (Preview)

Date published: 2021-05-21
Date modified: 2021-05-25

Legal Notice

Contents

Overview

Learn how to enable single sign-on (SSO) for Virtual Warehouse JDBC/ODBC connections in Cloudera Data Warehouse (CDW).

Virtual Warehouse JDBC/ODBC connections in CDW rely on the enterprise identity provider (IdP) you have set up in the User Management module of Management Console for SSO access. For information about configuring an IdP in CDP, see the link to "Configuring identity providers in CDP" at the bottom of this page. IdP providers that are compliant with Security Assertion Markup Language (SAML 2.0) can be used.

When CDW components use SSO, the connection URL includes the `auth=browser` clause. These connection URLs can be copied to your system for Hive or Impala Virtual Warehouses by selecting **Copy JDBC URL** in the options menu in the upper right-corner of the Virtual Warehouse tile in the CDW UI. If you are using ODBC to connect, ODBC connector (driver) documentation explains how to make the ODBC connection.

For example, the JDBC URL for an Impala Virtual Warehouse that is configured to use SSO in CDW might look like the following where the `auth=browser` clause is bolded:

Impala Virtual Warehouse configured to use SSO JDBC URL:

```
jdbc:impala://coordinator-test.env-mxxmcn.dw.xcu2-8y8x.dev.cldr.work:
443/default;
AuthMech=12;transportMode=http;httpPath=cliservice;ssl=1;auth=browser
;
```

In comparison, here is the JDBC URL for an Impala Virtual Warehouse without SSO enabled:

Impala Virtual Warehouse where SSO is not enabled (LDAP is used) JDBC URL:

```
jdbc:impala://coordinator-impala-without-sso.env-9qls4k.dw.xcu2-8y8x.
dev.cldr.work:443/default;AuthMech=3;transportMode=http;httpPath=clis
ervice;ssl=1;UID=csso_userID;
PWD=PASSWORD
```

The `auth=browser` clause is replaced with the user ID and password (`UID=csso_userID;PWD=PASSWORD`). When SSO is not enabled, LDAP is used.

> **Note:** You can configure an Impala Virtual Warehouse to use SSO and then connections that use the JDBC URL or the ODBC connection string are SSO-enabled connections. However, connections to the warehouse using the Impala shell will still use LDAP.

The following shows an equivalent comparison of JDBC URLs for a Hive Virtual Warehouse:

Hive Virtual Warehouse configured to use SSO JDBC URL:

```
jdbc:hive2://hs2-test.env-9qls4k.dw.xcu2-8y8x.dev.cldr.work/default;transportMode=http;httpPath=cliservice;ssl=true;auth=browser;
```

Hive Virtual Warehouse where SSO is not enabled (LDAP is used) JDBC URL:

```
jdbc:hive2://hs2-test-hive-without-sso.env-9qls4k.dw.xcu2-8y8x.dev.cldr.work/default;
transportMode=http;httpPath=cliservice;ssl=true;retries=3;
```

In the above case of the Hive Virtual Warehouse where SSO is not enabled, the `auth=browser` clause of the JDBC URL is replaced with `retries=3`.

> **Note**: If you have configured a Hive Virtual Warehouse for SSO-enabled connections, all connections are authenticated with the IdP provider that is configured in Management Console.

Authentication behavior when SSO is enabled

When you enable a Virtual Warehouse to use SSO in Cloudera Data Warehouse (CDW), authentication is performed via an interaction between your browser and your IdP provider.

Logging in the first time after configuring SSO/already logged in to CDW with your IdP credentials:

The following steps occur the first time you log into the Virtual Warehouse with your third-party BI tool after you have configured your Virtual Warehouse to use SSO. These steps occur when you are already logged in to the CDW experience with your IdP provider credentials:

1. In your BI tool, initiate a connection to a Virtual Warehouse that you have configured to use SSO.
2. A browser window appears with the following success message:
   "Successfully authenticated. You may close this window."
3. Close the browser window and use the third-party BI tool as needed.

Logging in subsequent times after SSO configuration/without first being logged in to CDW with your IdP credentials:

These steps occur after you have configured your Virtual Warehouse to use SSO, but you have not already logged in to CDW with your IdP credentials:

1. In your BI tool, initiate a connection to a Virtual Warehouse that you have configured to use SSO.
2. A browser window appears, which displays the login dialog for your IdP provider, offering you the ability to enter your login credentials, and submit them to your IdP provider.
3. A successful authentication message appears. You can close the browser window and resume work with your third-party BI tool.

Enable new Virtual Warehouses to use SSO

You can enable both Hive or Impala Virtual Warehouses to use single sign-on (SSO) by selecting this option when you first create them in Cloudera Data Warehouse (CDW). To use the SSO feature, the Virtual Warehouse must also be configured to use User Groups.

Before you begin

- Make sure you have configured an IdP in the User Management module of Management Console before you enable SSO for a Virtual Warehouse. For details, see "Configuring identity providers in CDP" in the Management Console documentation set.
- Determine the User Group, which has been set up in the User Management module of Management Console, that contains the users you want to have access to this Virtual Warehouse. For details, see "Managing groups in CDP."
- Make sure you have identified or created the following components in CDW:
  - An activated CDP environment. For details, see "Activating AWS environments" or "Activating Azure environments."
  - A Database Catalog. For details, see "Adding a Database Catalog."

Steps

1. Log in to the CDP web interface and navigate to the Data Warehouse service.
2. In the Data Warehouse service, click **Virtual Warehouses** in the left navigation panel.
3. On the Virtual Warehouses page, click **Add New**.
4. In the New Virtual Warehouse dialog box, specify the following:
   - **Name**
   - **Type** (Hive or Impala)
   - **Database Catalog** it queries
   - **Enable SSO** to use the IdP provider you configured in Management Console.
   - **User Groups** that can access endpoints
   - (Optional) **Tagging** keys and values
   - **Size**
5. After selecting the **Size** for the Virtual Warehouse, you can configure auto-scaling thresholds. For more information about tuning auto-scaling, see the links in the "Related information" section at the bottom of this page.
6. Click **Create** to create the new Virtual Warehouse.

Next steps

After you have configured a Virtual Warehouse to use SSO authentication, you must configure an ODBC or JDBC connection to the Virtual Warehouse with your third-party BI tool. For details, see Enable 3rd-party BI tools to use JDBC to connect to Virtual Warehouses or Enable 3rd-party BI tools to use ODBC to connect to Virtual Warehouses.

Enable third-party BI tools to use JDBC to connect to an SSO-enabled Virtual Warehouse

After you create an SSO-enabled Virtual Warehouse in CDW, you must configure your third-party BI tool to connect to the Virtual Warehouse. Learn how to configure your BI tool to use a JDBC connection.

Before you begin

- For JDBC connections to your Virtual Warehouse, download the JDBC connector (driver) from the Cloudera public website:
    a. Go to Hive JDBC Connectors for Hive Virtual Warehouses.
    b. Go to Impala JDBC Connectors for Impala Virtual Warehouses.
- Then select the following options:
    a. **Version** of the Hive or the Impala JDBC connector. Choose the latest version number.
    b. **SELECT AN OS** to select the operating system for which you are using the connector.
    c. Click **GET IT NOW** to download the JDBC connector.
  After you download the JDBC connector, also download its associated documentation. This documentation provides instructions for installing the JDBC connector and configuring a connection in your third-party BI tool. The documentation is linked to at the bottom of the download page for the Hive or the Impala JDBC connector.

- Make sure you are logged in to CDP and then navigate to CDW.
- Open your third-party BI tool and navigate to the location in the tool where you can configure a JDBC connection.

Steps

1. In your third-party BI tool, provide the path to the JDBC connector that you downloaded in the Before you begin section as required by the tool.

2. In CDW, click **Virtual Warehouses** in the left navigation menu, and click the name of the Virtual Warehouse you have configured to use SSO.

3. On the Virtual Warehouse details page, click the **Actions** menu in the upper right corner of the page, and select **Copy JDBC URL**, which copies the JDBC URL to your system's clipboard. This option is the same for both Impala and Hive Virtual Warehouses.

4. In your third-party BI tool, provide the JDBC URL as required by the tool, and initiate a connection to the Virtual Warehouse.

Enable third-party BI tools to use ODBC to connect to an SSO-enabled Virtual Warehouse

After you create an SSO-enabled Virtual Warehouse in CDW, you must configure your third-party BI tool to connect to the Virtual Warehouse. Learn how to configure your BI tool to use an ODBC connection.

Before you begin

- For ODBC connections to your Virtual Warehouse, download the ODBC connector (driver) from the Cloudera public website:
  a. Go to Hive ODBC Connectors for Hive Virtual Warehouses.
  b. Go to Impala ODBC Connectors for Impala Virtual Warehouses.
- Then select the following options:
  a. **Version** of the Hive or the Impala ODBC connector. Choose the latest version number.
  b. **SELECT AN OS** to select the operating system for which you are using the connector.
  c. Click **GET IT NOW** to download the ODBC connector.
  After you download the ODBC connector, also download its associated documentation. This documentation provides instructions for installing the ODBC connector and configuring a connection in your third-party BI tool. The documentation is linked to at the bottom of the download page for the Hive or the Impala ODBC connector.

- Make sure you are logged in to CDP and then navigate to CDW.

- Open your third-party BI tool and navigate to the location in the tool where you can configure an ODBC connection.

Steps

1. In your third-party BI tool, provide the path to the ODBC connector on your local system as required by the tool.

2. Refer to the documentation for your connector to specify the connection string in your 3rd-party tool, and initiate a connection to the Virtual Warehouse.

Related information
Configuring identity providers in CDP
Tuning Hive Virtual Warehouses
Tuning Impala Virtual Warehouses
Auto-scaling Virtual Warehouses
Custom tags in AWS environments
Custom tags in Azure environments