

CLOUDERA

Educational Services

Introduction to Cloudera Data Warehouse Self-Service Analytics in the Cloud with CDP: Hands-On Exercises

Table of Contents

How to Access CDW	1
CDW Orientation	2
Creating a Database Catalog	3
Creating a Virtual Warehouse	4
Accessing Query Editors	5
Setting Workload Password	6
Querying with Hue	7
Querying with DAS	8
Auto-Scaling	9
Additional Tuning for Hive	10
Additional Tuning for Impala	11
Using Impala Shell	12
Using Third-Party Tools	13

How to Access CDW

1. Access your CDP home page.
-

2. Click on the Data Warehouse icon to access CDW.
-

This is the end of the exercise.

CDW Orientation

Get accustomed to the CDW pages using the following suggestions.

1. Explore the sidebar: Click the grid to show the list of CDP applications, collapse and expand the sidebar, and view the Help.

2. Find the Overview page and expand the Environments section.

3. Click on an entity and look for the entities that are associated with it (note the Database Catalog and Virtual Warehouses).

4. Filter the Database Catalog or Virtual Warehouses section to show only a few, including the Database Catalog or Virtual Warehouse you just noted.

5. Go to the Database Catalogs page and find the same Database Catalog that you just noted, using the filter or page navigation if necessary.

6. Go to the Virtual Warehouses page and find the same Virtual Warehouses you just noted, using the filter or page navigation if necessary.

This is the end of the exercise.

Creating a Database Catalog

1. Create a Database Catalog for testing.
-

This is the end of the exercise.

Creating a Virtual Warehouse

1. Create a Virtual Warehouse associated with your test Database Catalog.
 - Use Hive.
 - Use the smallest size.

2. Suspend it and the test Database Catalog after it's been created.

This is the end of the exercise.

Accessing Query Editors

1. Open a query editor.

- You don't have to query with it (that will happen in another section), just be sure you can open it.
- If necessary, continue to next video to set up your workload password.

This is the end of the exercise.

Setting Workload Password

1. Create your workload password:
 - a. Enter CDW.
 - b. Click your name in the bottom left, then **Profile**.
 - c. Click the **Set Workload Password** link.
 - d. Fill in the necessary information and click the **Set Workload Password** button.

2. Test it by accessing Hue or DAS.

This is the end of the exercise.

Querying with Hue

1. In Hue, complete a simple `SELECT *` query on a table you can access.
-

This is the end of the exercise.

Querying with DAS

1. In DAS, complete a simple `SELECT *` query on a table you can access.
-

This is the end of the exercise.

Auto-Scaling

1. Change the auto-scale settings:
 - For a new Virtual Warehouse, on creation
 - For an existing Virtual Warehouse
-

This is the end of the exercise.

Additional Tuning for Hive

1. Change the Headroom/Wait Time settings for a new or existing Hive Virtual Warehouse.
-

This is the end of the exercise.

Additional Tuning for Impala

1. Change the Autoscale Mode setting for a new or existing Impala Virtual Warehouse.
-

This is the end of the exercise.

Using Impala Shell

1. Install Impala Shell using the following steps, unless you are using a cluster node.

a. Open a terminal window.

b. Run this command:

```
$ pip install impala-shell
```

c. Verify it was installed using this command:

```
$ impala-shell --help
```

2. Use Impala Shell to query a table.

a. Copy the shell command from your Virtual Warehouse in CDW.

b. Run that command in a terminal window.

c. Run a SQL command to query a table. Use `SHOW DATABASES;` or `SHOW TABLES;` if needed, or to verify the connection first.

This is the end of the exercise.

Using Third-Party Tools

1. Connect a business intelligence (BI) application to Hive:

a. Download a driver using one of these steps:

- Go to cloudera.com and use the product downloads to download an ODBC driver for Hive.
- Use the options menu for the Virtual Warehouse to download the JDBC Jar file.

Note: You will need to consult your BI documentation to locate where to place the driver.

b. Download a BI if necessary.

c. Complete the connection:

- Use the options menu for the Virtual Warehouse to copy the JDBC URL.
- Paste the URL into a text editor and copy just the server portion of the URL.
- Find the settings for connecting to a server within your BI. Use the following:

Setting	Value
Server	The copied server address
Port	443
Type	HiveServer2
Authentication	Username and Password (use your credentials)
Transport	HTTP
Path	cliserver

d. Verify the connection is complete by finding a table in your Database Catalog.

This is the end of the exercise.