Cloudera Runtime 1.5.5

Administering Hue

Date published: 2020-08-10
Date modified: 2025-06-06

CLOUDERA

https://docs.cloudera.com/

Legal Notice

Contents

Hue configurations in Cloudera Data Warehouse

You can customize the settings for Hue at a Vitrual Warehouse level by specifying the configurations in the hue-safety-valve field in the Cloudera Data Warehouse UI.

The following table lists the safety valve parameters supported by Hue in Cloudera Data Warehouse:

Parameter	Description

`[notebook]`
`  [[interpreters]]`
`  [[[hive]]]`
`  name=Hive`
`  interface=hiveserver2`
`  [[[hplsql]]]`
`  name=Hplsql`
`  interface=hiveserver2`
	Used to activate and enable switching between Hive and HPL/SQL interpreters for your queries.
	Note: Hue enables you to switch between Hive and HPL/SQL interpreters. By default, the Hive interpreter is enabled when you create a Hive Virtual Warehouse. To enable the HPL/SQL interpreter, you must update the configuration in the hue-safety-valve field in your Hive Virtual Warehouse. However, updating hue-safety-valve overrides the default configuration. Therefore, to use both Hive and HPL/SQL interpreters, you must enable both by updating the configuration in the hue-safety-valve field.

`[[desktop]]` `app_blacklist`	Used to add or remove applications, such as the File Browser, Impala, Hive, Oozie, and so on from the blocked list of applications.

`[desktop]`
`  [[session]]`
`  ttl=[***NUMBER-OF-SECONDS***]`
	Used to configure the duration of a user session. The ttl property determines the length of time that the cookie with the user's session ID lives before expiring. After the ttl setting is reached, the user's session expires whether it is active or not.
	The default setting for ttl is 1,209,600 seconds, which equals two weeks.

`[jobbrowser]`
`[[query_store]]`
`server_url=[***QUERY-PROCESSOR-URL***]`
	Used to display the **Queries** tab for Hive and Impala on the **Job Browser** page. This configuration is enabled by default and is not exposed in the Hue safety valve.
	However, to hide the **Queries** tab, you can override the configuration by adding the following lines in Virtual Warehouse ⋮ Edit CONFIGURATIONS Hue hue-safety-valve :
	`[jobbrowser]`
	`[[query_store]]`
	`is_enabled=false`
	To enable the **Queries** tab, set is_enabled to true.

Hue supported browsers

Hue works with the two most recent LTS (long term support) or ESR (extended support release) browsers. Cookies and JavaScript must be enabled.

The lists the minimum tested versions of the most common browsers:

- Chrome: (Version history)
- Firefox: (Version history)
- Safari (Mac only): Version history
- Microsoft Edge: (Version history)

Hue can display in other browsers and in older versions of the common browsers, but you might not have access to all features.

Enabling cache-control HTTP headers when using Hue

You can enable Hue to use HTTP headers such as Cache-Control, Pragma, and Expires to ensure that your browser always uses the fresh version of the resource. You must set the value of the custom_cache_control property to true in Hue's Advanced Configuration Snippet.

About this task

Note: Enabling the use of these cache-control HTTP headers can affect performance because your browser tries to fetch the latest resources instead of using cached data. To disable cache-control HTTP headers, set the value of the custom_cache_control property to false.

Procedure

1. Go to Clusters Hue Configuration and add the following line in the Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini field:

```
[desktop]
custom_cache_control=true
```

2. Click Save Changes.
3. Restart the Hue service.

Customizing the Hue web interface

You can customize the page logo and set the cache timeout value by configuring the parameters in the Virtual Warehouse which is running Hue.

Adding a custom banner in Hue

You can add a custom banner to the Hue web interface by adding your custom HTML to the hue-safety-valve configuration for your Virtual Warehouse.

Procedure

1. Log in to the Data Warehouse service as DWAdmin.
2. Go to your Virtual Warehouse tile and click ⋮ Edit .
3. Go to CONFIGURATIONS Hue , select hue-safety-valve from the Configuration files drop-down menu and add the following lines:

```
[desktop]
[[custom]]
banner_top_html=<H1>Your company's custom Hue Web UI banner</H1>
```

4. Click Apply Changes.

   The Virtual Warehouse goes into an "Updating" state. Wait for the update to complete, and then open or refresh Hue.

Changing the page logo in Hue

You can replace the Hue web interface logo with a custom log that is created with an SVG code. Add any type of logo you want, but your custom logo should be designed to fit into a 160 x 40 pixel space.

About this task

For example, here is the Hue logo shown in the following image:

You can change this Hue logo by adding the appropriate SVG code to the logo_svg property under the [desktop] [[custom]] section in the hue_safety_valve configuration parameter in Cloudera Data Warehouse.

Procedure

1. Log in to the Cloudera Data Warehouse service as an administrator.
2. Go to the Virtual Warehouses Edit CONFIGURATIONS Hue and select hue-safety-valve from the Configuration files drop-down list.
3. Add the custom logo SVG code in the [desktop] [[custom]] section as shown in the following example:

```
[desktop]
[[custom]]
logo_svg='[***SVG-CODE-FOR-CUSTOM-LOGO***]'
```

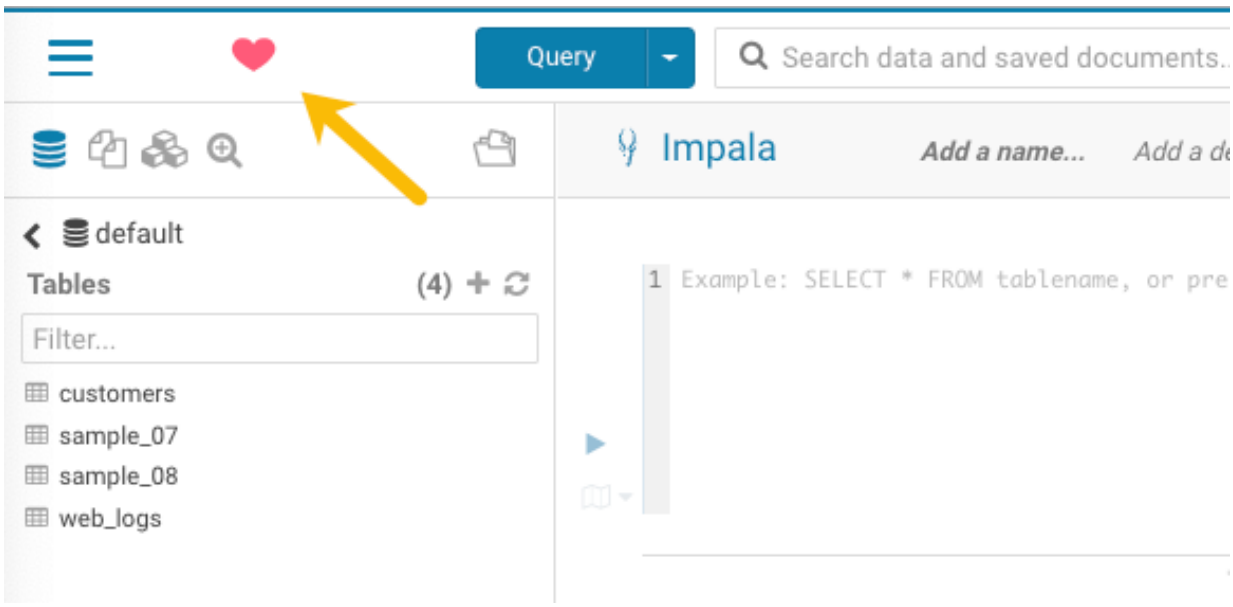
For example, the following SVG code replaces the Hue logo with a red heart:

```
[desktop]
[[custom]]
logo_svg='<g><path stroke="null" id="svg_1" d="m44.41215,11.43463c-4.05
017,-10.71473
-17.19753,-5.90773 -18.41353,-0.5567c-1.672,-5.70253 -14.497,-9.95663
-18.411,0.5643c-4.35797,11.71793 16.891,22.23443 18.41163,23.95773c1.518
1,-1.36927 22.7696,-12.43803
18.4129,-23.96533z" fill="#ffffff"/> <path stroke="null" id="svg_2"
d="m98.41246,10.43463c-4.05016,-10.71473 -17.19753,-5.90773 -18.41353,-
0.5567c-1.672,-5.70253
```

```
-14.497,-9.95663 -18.411,0.5643c-4.35796,11.71793 16.891,22.23443 18.4116
4,23.95773c1.5181,-1.36927
22.76959,-12.43803 18.41289,-23.96533z" fill="#FF5A79"/> <path stroke="nu
ll" id="svg_3"
d="m154.41215,11.43463c-4.05016,-10.71473 -17.19753,-5.90773 -18.41353,-0
.5567c-1.672,-5.70253
-14.497,-9.95663 -18.411,0.5643c-4.35796,11.71793 16.891,22.23443 18.41164
,23.95773c1.5181,-1.36927 22.76959,-12.43803 18.41289,-23.96533z" fill="
#ffffff"/> </g>'
```

4. Click APPLY.
5. Restart the Virtual Warehouse.
6. Verify your changes by opening Hue.

   If you added the sample SVG code that defines a red heart as the logo, then your Hue web interface looks as shown in the following image:



Related Information

Scalable Vector Graphics

Adding a splash screen in Hue

You can add a custom splash screen to the Hue web interface by adding your custom HTML to the hue-safety-valve configuration for your Virtual Warehouse.

Procedure

1. Log in to the Cloudera Data Warehouse as DWAdmin.
2. Go to your Virtual Warehouse tile and click ⋮ Edit .
3. Go to CONFIGURATIONS Hue , select hue-safety-valve from the Configuration files drop-down menu and add the following lines:

```
[desktop]
[[custom]]
```

```
login_splash_html=[***CUSTOM-HTML***]
```

```
[desktop]
[[custom]]
login_splash_html=<h1>Hue, the next-gen SQL Assistant</h1>
```

4. Click Apply Changes.

   The Virtual Warehouse goes into an "Updating" state. Wait for the update to complete, and then open or refresh Hue.

Setting the cache timeout

Enable Hue UI caching by setting a timeout value in milliseconds. The default is 10 days or 864000000 milliseconds. Set the timeout to 0 to disable caching.

About this task

When you browse tables using the left assist panel or run queries, Hue caches this information for fetching information faster and query autocompletion. You can configure the time for which you want to Hue to cache this information by setting the value of the cacheable_ttl property under the [desktop][[custom]] section in the hue_safety_valve configuration property in Cloudera Data Warehouse.

Procedure

1.
2. Log in to the Cloudera Data Warehouse service as an administrator.
3. Go to the Virtual Warehouses ⋮ Edit CONFIGURATIONS Hue and select hue-safety-valve from the Configuration files dropdown menu.
4. Add the following parameters with the cache timeout value to the hue_safety_valve configuration parameter:

   ```
   [desktop]
   [[custom]]
   cacheable_ttl=[***VALUE-IN-MILLISECONDS***]
   ```

   For example, the following configuration sets the cache timeout to the default value of 86400000 milliseconds:

   ```
   [desktop]
   [[custom]]
   cacheable_ttl=86400000
   ```

5. Click APPLY.
6. Restart the Virtual Warehouse.

Enabling or disabling anonymous usage date collection

Hue tracks anonymized pages and application versions to gather information about application usage levels. The data collected does not include hostnames or IDs. For example, the data collected has the format /2.3.0/pig or /2.5.0/beeswax/execute.

About this task

To enable or disable anonymous usage data collection:

Procedure

1. In the Cloudera Manager Admin Console, select Clusters Hue Configuration to navigate to the configuration page for Hue.
2. In the Search text box, type usage to locate the Enable Usage Data Collection check box:
   - To enable anonymous data collection, check the box, which is the default setting.
   - To disable anonymous data collection, clear the check box.
3. Enter a Reason for change..., and then click Save Changes at the bottom of the page to save the configuration change.
4. Refresh the browser page and click the restart icon  at the top of the page so the new configuration changes can be read by the server and the new data collection setting takes effect.

Disabling the share option in Hue

Hue allows you to share documents, queries, and workflows with other users, either for viewing only or viewing and modifying in any Hue instances across all Virtual Warehouses within a Database Catalog. Sharing is enabled by default in the Hue UI. For added privacy and control, you can disable sharing by setting the enable_sharing property to false in the hue-safety-valve.

About this task

The sharing option is always available to the admin users. To disable the share option:

Procedure

1. Log in to the Cloudera Data Warehouse service as an administrator.
2. Go to the Virtual Warehouses Edit CONFIGURATIONS Hue and select hue-safety-valve from the Configuration files drop-down list.
3. Add the following lines in the hue-safety-valve:

```
[desktop]
enable_sharing=false
```

4. Click APPLY.
5. Restart the Virtual Warehouse.

Adding Query Processor Administrator users and groups in Cloudera Data Warehouse

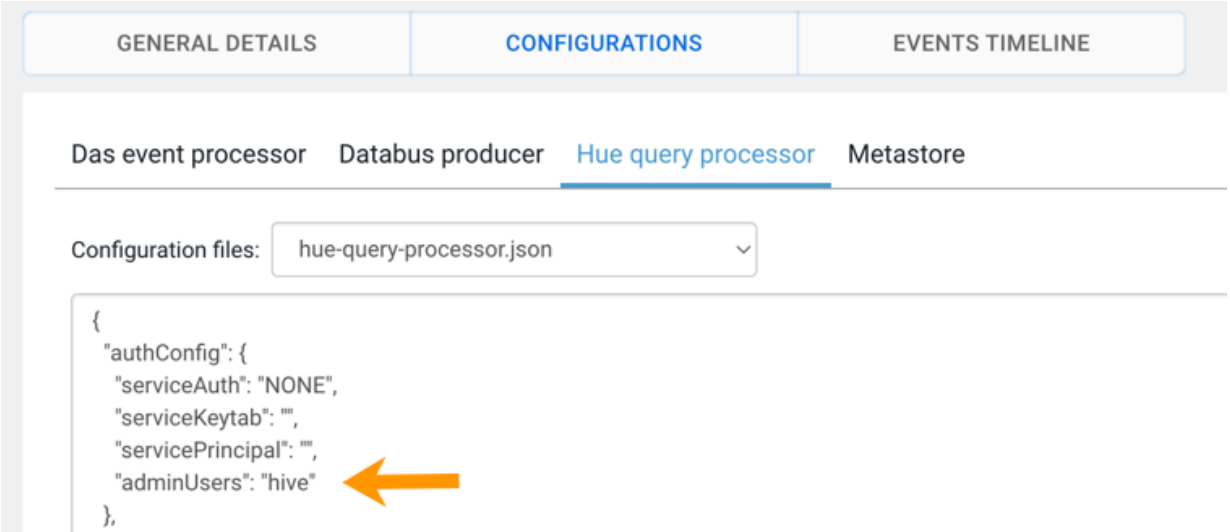
The Query Processor Administrators have special privileges that enable them to view and monitor queries from all users, including the ones that were submitted from query interfaces, such as Beeline, Hive Warehouse Connector (HWC), Tableau, Impala-shell, Impyla, and other JDBC/ODBC clients.

Before you begin

Make sure that the Virtual Warehouse to which you want to add the Hue Query Processor Administrators users is in the stopped state.

Procedure

1. Log in to the Cloudera Data Warehouse web interface as a DWAdmin.
2. Click ⋮ Edit on the Database Catalog for which you want to add Hue Query Processor Administrators users.
3. On the **Database Catalog** details page, click CONFIGURATIONS Hue query processor and select hue-query-processor.json from the Configuration files drop-down menu, as shown in the following image:



4. In the “authConfig” section, add the list of users to the “adminUsers” key.
   For example: "adminUsers": "hive, [***USER-1***], [***USER-2***]"
   You can also add a list of admin groups as follows:

```
"adminGroups": "admin-group, [***GROUP-1***], [***GROUP-2***]"
```

5. Click Apply.
   The Hue service will be unavailable for approximately 5 minutes to make the update.

Ways to clean up old queries from the Query Processor tables

Learn how to schedule a query clean-up and how to use the API to manually clean up queries from the following Query Processor tables: vertex_info, dag_details, dag_info, query_details, hive_query, tez_app_info.

Scheduling query clean-up

Both Hive and Impala queries are retained in the backend database for 30 days by default, after which they are cleaned up. You can change the clean-up interval from the Database Catalog configurations. Go to Database Catalogs ⋮ Edit CONFIGURATIONS Hive query processor and select the hue-query-processor.json from the Configuration files drop-down menu.

Add the following line under the dasConf section and specify the time interval in seconds:

```
 "dasConf": {
    "hue.query-processor.event-pipeline.cleanup-interval-secs": "[***TIME-
INTERVAL-IN-SECONDS***]",
    “hue.query-processor.event-pipeline.cleanup.cron.expression” : “[***CR
ON-EXPRESSION***]”
```

```
  },
```

For example:

```
“hue.query-processor.event-pipeline.cleanup.cron.expression” : “0 0 2 * * ?”
,
“hue.query-processor.event-pipeline.cleanup-interval-secs” : “2592000”
```

Manually cleaning up queries using an API

The ability to clean up queries manually in addition to the scheduled clean-up routines is useful when you have a high load of queries in a particular week that are hogging resources that you must free up. The API also runs a VACUUM command on the Query Processor table to reclaim storage that is occupied by dead tuples.

You can send an API request using tools such as cURL or Postman.

API format: `[***X-Do-As:[COMPONENT/PROCESS USER ID]***][***QUERY-PROCESSOR-ADDRESS***]/api/admin/cleanup/[***EPOCH-TIME***]`

Where,

- [***QUERY-PROCESSOR-ADDRESS***] is the query processor host address
- [***EPOCH-TIME***] is the Unix epoch time in seconds

Queries that were run before the specified epoch time are purged.

For example:

```
curl -k --negotiate -u : -H "X-Do-As: hive" "https://machine1.company.com:30
700/api/admin/cleanup/1670006742"
```

Downloading debug bundles

The debug bundle is a ZIP file that contains the query details in JSON format and an error-reports.json file, which is created only if an error occurs while the query is run.

About this task

Note: This feature is available only for Hive queries.

If Tez is used to run a query, then the debug bundle also contains DAG and Tez JSON files, as shown in the following image:

Procedure

1. Go to the Cloudera Data Warehouse web interface and open Hue from your Virtual Warehouse.
2. Click Jobs from the left assist panel.
   The **Job Browser** page is displayed.
3. Click Queries.
   The Hive queries that were run are displayed.
4. Select a query for which you want to download the debug bundle.
5. Click Download and save the ZIP file on your computer.
   The filename is in the following format:

```
hive_[***HIVE-QUERY-ID***]_[***USER-ID***]_[***UNIQUE-INDEX***]
```

Enabling Spark 3 engine in Hue

Before you begin

Note: Livy v0.6 supports Python versions upto Python 3.7. If you install Python 3.8 or higher, then you may see the following error: TypeError: required field "type_ignores" missing from Module.

Procedure

1. Log in to Cloudera Manager as an Administrator.
2. Go to Clusters HDFS Configuration and add the following lines in the Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml field:

```
<property>
<name>hadoop.proxyuser.hue.groups</name>
<value>*</value>
</property>
<property>
<name>hadoop.proxyuser.hue.hosts</name>
<value>*</value>
</property>
<property>
<name>hadoop.proxyuser.spark.groups</name>
<value>*</value>
</property>
<property>
<name>hadoop.proxyuser.spark.hosts</name>
<value>*</value>
</property>
<property>
<name>hadoop.proxyuser.livy.groups</name>
<value>*</value>
</property>
<property>
<name>hadoop.proxyuser.livy.hosts</name>
<value>*</value>
</property>
```

3. Click Save Changes.
4. Go to Clusters Livy for Spark 3 service Configuration and add the following configurations:
   a) Add the hue user in the Admin Users (livy.superusers) field.
   b) Go to the HMS Service field and select Hive.
   c) Click Save Changes.
5. Go to Clusters SPARK_ON_YARN Configuration Admin Users , add hue to the list of admin users (spark.history.ui.admin.acls) and click Save Changes.
6. Go to Clusters SPARK Configuration Admin Users , add hue to the list of admin users (spark.history.ui.admin.acls) and click Save Changes.
7. Go to Clusters SPARK 3 Configuration Admin Users , add hue to the list of admin users (spark.history.ui.admin.acls) and click Save Changes.
8. Go to Clusters Hue Configuration and enter the following lines in the Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini field and click Save Changes:

```
[desktop]
app_blacklist=zookeeper, pig #custom list of blocked apps
[spark]
#This is not a thrift server port
#If this TLS/SSL is enabled then check to see whether the livy url is on h
ttps or http and modify the url accordingly.
livy_server_url=http(s)://[***LIVY-FOR-SPARK3-SERVER-HOST***]:[***LIVY-
FOR-SPARK3-SERVER-PORT***]
ssl_cert_ca_verify=false
security_enabled=true
[notebook]
[[interpreters]]
[[[sparksql]]]
name=Spark SQL
```

```
interface=livy
```

Attention:

- Ensure that the Spark application is not on the blocked list.
- Set ssl_cert_ca_verify=false if an SSL certificate is not present in Hue's truststore. ssl_cert_ca_verify=true if your environment is a secured environment.
- Set security_enabled=true for Kerberized clusters.

9. Restart the affected services.

Results

You can now select the “Spark SQL” dialect on the Hue editor and run Spark queries from Hue.

Note: Starting a Livy session can take 30-45 seconds. If you cancel a running Spark SQL query and rerun it, the Livy session expires, and a new session is created.

Managing file extensions for Hue uploads

You can allow and restrict file extensions across all configured file systems in Hue to control file uploads. For instance, you can allow .csv files while blocking .exe files.

Procedure

1. Log in to Cloudera Manager Clusters Hue service Configuration .
2. In the search, enter hue_safety_valve.ini to locate the Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini field.
3. Set the configuration in the Hue Service Advanced Configuration Snippet (Safety Valve) for the hue_safety_valve.ini field as follows:

```
[filebrowser]
restrict_file_extensions = .sh, .exe #Specify file extensions that are not
 allowed, separated by commas.
allow_file_extensions = .tsv, .csv  #Specify file extensions that are al
lowed, separated by commas.
```

Note:

- By default, no file extensions are restricted during file uploads.
- Files are allowed only if they match the allowed file extension type and are not listed in the restricted file extension types. Otherwise, an error message is displayed.

4. Click Save Changes.
5. Restart the Hue service.

Enabling browsing Ozone from Hue on Cloudera Data Warehouse

Hue can read and write files on the Ozone filesystem, similar to S3 or ADLS. To access Ozone from Hue, you must add the additional configurations in the in the hue-safety-valve within the Virtual Warehouse configurations.

About this task

You can perform this task on any Hue instances of any environment in which you want to enable the Ozone File Browser. If multiple Hue instances exist within the same cluster, completing this task on one instance per environment is sufficient.

Before you begin

- Ensure that the Ozone HttpFS Gateway role is running in a healthy state.
- Go to Cloudera Manager Ozone Configuration and add the following entries in the HttpFS Gateway Advanced Configuration Snippet (Safety Valve) for ozone-conf/httpfs-site.xml field:

Field name	Value
Name	httpfs.proxyuser.[***PRINCIPAL-NAME***].hosts
Value	*
Name	httpfs.proxyuser.[***PRINCIPAL-NAME***].groups
Value	*

  Replace the [***PRINCIPAL-NAME***] with the actual Kerberos principal name. The **hive** principal is the default principal required for communication between Ozone and Hue. If this principal name is changed during installation (e.g., to a custom principal for the Hive service), use that modified principal name here instead.
- Obtain the following values from the Ozone service. You need it to construct the fs_defaultfs and webhdfs_url URLs:
  - HttpFS Gateway host name (Gateway node)
  - Ozone HttpFS Gateway TTP Web UI Port (ozone.httpfs.http-port)

    The default port is 9778. Ensure that the port used by the Ozone HttpFS Gateway (or any other configured port for that HttpFS instance) is accessible from the Hue node to the Gateway node where the Ozone HttpFS Gateway is installed.
  - Ozone Service ID (ozone.service.id).

Note: If you upgrade to version 1.5.5 SP1, the following manual Hue safety valve configuration for Ozone is no longer required. The necessary configuration settings are now applied automatically, and Ozone is enabled by default.

Procedure

1. Log in to the Cloudera Data Warehouse service as DWAdmin.
2. Go to Virtual Warehouse ⋮ Edit CONFIGURATIONS Hue and select hue-safety-valve from the Configuration files drop-down menu. Replace the existing configuration with the following lines.

```
[desktop]
rest_conn_timeout=300
enable_chunked_file_uploader=false
[[task_server_v2]]
enabled=false
[[ozone]]
[[[default]]]
fs_defaultfs=ofs://[***SERVICE-ID***] \\The ozone.service.id value
webhdfs_url=https://[***OZONE-HTTPFS-HOST***]:[***OZONE-HTTPFS-PORT***]/w
ebhdfs/v1
ssl_cert_ca_verify=true
security_enabled=true
[hadoop]
```

```
upload_chunk_size=2147483648
```

Configurations	Descriptions
`rest_conn_timeout=300`	This optional configuration extends the default data push timeout in Ozone from 120 seconds to 300 seconds. This can enhance performance when transferring larger data chunks to Ozone, especially over slower network connections.

`enable_chunked_file_uploader=false`
`[[task_server_v2]]`
`enabled=false`
	This configuration is necessary because Ozone does not support chunked uploads or the `/append` HttpFS API. Therefore, files must be uploaded as a single, complete chunk into Ozone

`[[ozone]]`
`[[[default]]]`
`fs_defaultfs=ofs://[***SERVICE-ID***] \\The ozone.service.id value`
`webhdfs_url=https://[***OZONE-HTTPFS-HOST***]:[***OZONE-HTTPFS-PORT***]/webhdfs/v1`
`ssl_cert_ca_verify=true`
`security_enabled=true`
	This configuration enables you to browse objects within Ozone DB buckets from Hue.

`[hadoop]` `upload_chunk_size=2147483648`	This configuration increases the default file upload chunk size from 64 MB to 2 GB, which is the maximum supported by the Django code used by Hue when uploading files.

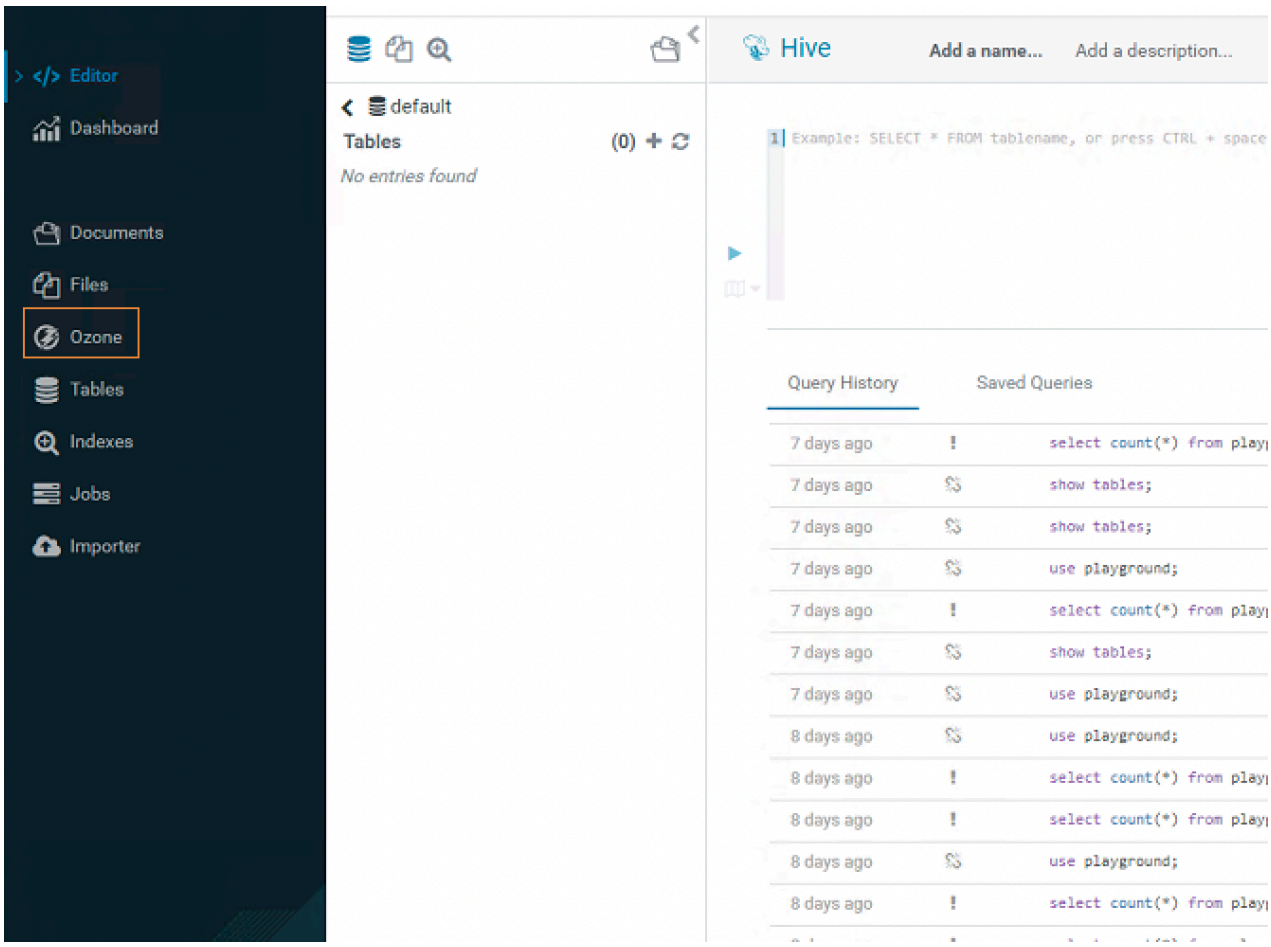
Note:

- While you can create new volumes/buckets in Ozone through the Hue interface, managing or erasing them requires the Ozone CLI.
- If the cluster is not secured, set the webhdfs_url property to http.
- If Ozone is not configured in HA mode, then specify the URL for the Ozone manager in the fs_defaultfs property.

3. Click Apply Changes.
4. Log in to Hue as an Administrator on any one instance within the environment.
5. Click your username in the lower-left corner of the interface, and select Administer Users.
6. Navigate to the Groups tab, select the default group, and ensure filebrowser.ofs_access:Access to OFS from filebrowser and filepicker permission is selected.
7. Click Update group to save the changes.
8. Restart the Hue service.

Results

After configuring the Hue safety valve and restarting the Virtual Warehouse, you may see that the Ozone file browser may take 10 - 15 minutes to display on the Hue web interface, as shown in the following image.


Editor
Dashboard
Documents
Files
Ozone
Tables
Indexes
Jobs
Importer
default
Tables (0)
No entries found
Hive
Add a name...
Add a description...
1 Example: SELECT * FROM tablename, or press CTRL + space
Query History
Saved Queries
7 days ago select count(*) from play
7 days ago show tables;
7 days ago show tables;
7 days ago use playground;
7 days ago select count(*) from play
7 days ago show tables;
7 days ago use playground;
8 days ago use playground;
8 days ago select count(*) from play
8 days ago select count(*) from play
8 days ago use playground;
8 days ago select count(*) from play