Cloudera Data Warehouse Private Cloud 1.5.4

Data Warehouse Release Notes

Date published: 2020-08-17
Date modified: 2024-07-31

CLOUDERA

Legal Notice

Contents

What's new in Cloudera Data Warehouse Private Cloud

Review the new features in this release of Cloudera Data Warehouse (CDW) service on Private Cloud.

Hue supports natural language query processing (Preview)

Hue leverages the power of Large Language Models (LLM) to help you generate SQL queries from natural language prompts and also provides options to optimize, explain, and fix queries, promoting efficient and accurate practices for accessing and manipulating data. You can use several AI services and models such as OpenAI's GPT service, Amazon Bedrock, and Azure's OpenAI service to run the Hue SQL AI assistant.

- To learn more about the supported models and services, limitations, and what data is shared with the LLMs, see About the SQL AI Assistant in CDW.
- To set up and enable the SQL AI Assistant, see About setting up the SQL AI Assistant in CDW.
- To see how to generate, edit, explain, optimize, and fix queries, see Starting the SQL AI Assistant in Hue.

Improvements to custom pod configuration (now known as Resource Templates)

Several improvements and changes have been made to the custom pod configuration functionality starting with CDP Private Cloud Data Services 1.5.4. The custom pod configuration has been renamed to "Resource Templates".

- A new menu option Resource Templates has been added to the left navigation pane on the CDW web interface.
- You can now configure the allocation of Kubernetes resources to the pods for Hive, Data Visualization, and Database Catalog in addition to Impala.
- The Impala pod configuration feature is moved from the **Environment Details** page to the **Resource Templates** page.

For more information, see Resource templates for CDW Private Cloud pods.

Flexibility to enable and disable quota management for CDW entities

Earlier, you were required to enable the quota management option before activating an environment to use quota-managed resource pools for environments, Data Catalogs, Virtual Warehouses, and Data Visualization instances in CDW. Starting with CDP Private Cloud Data Services 1.5.4, you can enable or disable the quota management option at any time during the lifecycle of the CDW entities. To learn more about the behavioral aspects, see Quota Management in CDW Private Cloud.

A new option called --resource-pool (string) has been added to the CDP CLI commands for CDW Private Cloud using which you can specify the resource pool for the CDW entity (Virtual Warehouse, cluster/environment, Data Visualization instance).

Added support for authentication between CDW and HMS database using mTLS

CDW and the Hive MetaStore (HMS) database on the base cluster can mutually authenticate each other during the SSL/TLS handshake using mTLS for all supported backend databases (Oracle, MySQL, MariaDB, and Postgres). To set up mTLS, you must upload the database client certificate and the private key files in PEM format while activating an environment in CDW. See Enabling mTLS between the HMS database and CDW on Private Cloud.

Ability to enable active-passive configuration for HiveServer2 pods

CDW provides an option to enable active-passive configuration for HiveServer2 (HS2) pods in Private Cloud. By enabling this feature, two HS2 pods run simultaneously–one active and the other inactive. When one pod terminates, the inactive pod becomes active, providing High Availability. See HiveServer2 High Availability in CDW Private Cloud. The most likely cause of a pod's termination is node failure.

Improvements to backup and restore CDW

There are two ways to backup and restore the Data Warehouse service:

- Using Data Recovery Service (DRS)
- Using the CDW's CDP CLI cluster management commands

Both these methods use CDP CLI commands. The CDP Management Console also provides a graphical user interface to perform backup and restore operations. See DRS automatic backups and Using DRS with CDW.

To learn about the different backup and restore methods, objects and configurations that are included in the backup, deciding which method to use, see Backup and restore in CDW Private Cloud.

New advanced configuration option to backup Virtual Warehouse namespaces before an upgrade

By default, CDW backs up namespace-related data before starting the upgrade process using the Data Recovery Service (DRS). A new option has been added to disable the automatic backup process on the **Advanced Configuration** page in the CDW web interface.

CDW no longer has a dependency on YARN

Environment activation in CDW no longer depends on or fails if the YARN service is not installed on the CDP Base cluster.

Known issues and limitations in Cloudera Data Warehouse Private Cloud

Review the known issues and limitations that you might run into while using the Cloudera Data Warehouse (CDW) service in CDP Private Cloud Data Services.

General known issues

This topic describes the general service-wide known issues for Cloudera Data Warehouse (CDW) Private Cloud.

Known issues identified in 1.5.4

DWX-18558: The executor pods in Impala Virtual Warehouse do not update when you change it to use a different resource template

Suppose you created an Impala Virtual Warehouse with a certain resource template. If you apply a different resource template later having a different local storage size, the operation fails silently and the following pods are not updated: hiveserver 2, impala-coordinator, impala-executor and hue-backend. This happens because changing the storage size is not supported by Kubernetes.

None. Cloudera recommends that you avoid changing resource templates with different volume sizes and select the right size while creating the Virtual Warehouse.

Known issues identified in 1.5.3

DWX-17880: Hive Virtual Warehouse does not start if the bind user contains special characters

The Hive virtual warehouse may fail to start up if you have specified the following special characters in the LDAP bind credential password: $< > \& ' "$. This happens because the HiveServer2 (HS2) pod gets stuck in the CrashLoopBackOff state with the following error in its logs: error parsing conf file:/etc/hive/conf/hive-site.xml com.ctc.wstx.exc.WstxUnexpectedCharException: Unexpected character '&' (code 38) in content after '<' (malformed start element?). at [row,col,system-id]: [388,13,"file:/etc/hive/conf/hive-site.xml"].

1. Change the LDAP bind credentials in the Management Console. ensure that they do not contain the following unsupported special characters: < > & '   ".
2. Reactivate the environment in CDW.

Known issues identified in 1.5.1

DWX-15142 Character restriction on environment name when using FreeIPA server version 4.9.8 and higher

FreeIPA is supported as an authentication mechanism starting with the 1.5.1 release. If you are using FreeIPA version 4.9.8 and higher, then note that the host names are limited to 64 characters. Because the environment name is part of the host name, the environment name must not exceed 17 characters.

None.

Known issues identified in 1.5.0

DWX-18903: Service "postgres-service-default-warehouse" is invalid: spec.externalName error

You see the following error during the Database Catalog creation stage after activating the environment in CDW:

```
Service "postgres-service-default-warehouse" is invalid:
 spec.externalName
a lowercase RFC 1123 subdomain must consist of lower case
 alphanumeric characters, '-' or '.', and must start and end with
 an alphanumeric character (e.g. 'example.com', regex used for
 validation is '[a-z0-9]([-a-z0-9]*[a-z0-9])?(\.[a-z0-9]([-a-
z0-9]*[a-z0-9])?)*')
```

. This could happen because if the value of the Hive Metastore Database Host (hive_metastore_database_host) property on the base cluster is not specified in lowercase.

Go to Cloudera Manager Clusters Hive service Configuration and change the value specified in the Hive Metastore Database Host field to be in lowercase.

Known issues identified before 1.4.1

DWX-10403: Executor pods get stuck in pending state with a warning

In rare circumstances, when Impala or Hive executors start up either due to autoscaling or by manually restarting the executors, the pods may get stuck in a pending state with a warning such as "volume node affinity conflict". This happens due to a race condition in the storage class that provides local volumes.

Restart the pods so that they can be rescheduled on new nodes with enough resources.

DWX-8502: HMS health check does not check port 9083

The HMS health check script does not check the health of its service port 9083 and may provide incorrect health status.

None.

Upgrade-related known issues

This topic describes the upgrade-related known issues for Cloudera Data Warehouse (CDW) Private Cloud.

Known issues identified in 1.5.4

DWX-18447: Virtual Warehouses go into an erroneous state after upgrading the Control Plane

After upgrading the Control Plane to CDP Private Cloud Data Services 1.5.4, you may see that the Virtual Warehouse creation fails with the following error:

```
huefrontend-5888cc97b8-fxpx5 pod in compute-1716270252-fzfj
 namespace has an erroneous container, state: waiting, reason:
 RunContainerError, message: context deadline exceeded Error
 Code : undefined
```

When you run the `kubectl describe` command on the pod, you see the following warning:

```
Error: failed to create containerd task: failed to create shim
 task: OCI runtime create failed: runc create failed: unable to
 create new parent process: namespace path: lstat /proc/0/ns/ipc:
 no such file or directory: unknown
```

None. The Virtual Warehouse transitions back into the running or stopped state after the pod has initialized successfully.

Known issues on OpenShift cluster environments

This topic describes the OpenShift cluster environment known issues for Cloudera Data Warehouse Private Cloud.

Known issues identified in 1.5.4
DWX-18151: Hue backup and restore fails on OpenShift Container Platform (OCP)

The file permissions for the pgpass file are set incorrectly on mount, due to restricted Security Context Constraint on OCP. This causes failure in connecting to the Hue database, resulting in failure of the job.

None.

ECS cluster environments

This topic describes the Embedded Container Service (ECS) cluster environment known issues for Cloudera Data Warehouse Private Cloud.

Known issues identified in 1.5.4

No new known issues identified in 1.5.4.

Known issues identified before 1.4.1
BLESC-6074: Impala executor pods not running and queries from Hue fail

You may see the following error after submitting a query from Hue: Latest admission queue reason: Waiting for executors to start. Only DDL queries and queries scheduled only on the coordinator (either NUM_NODES set to 1 or when small query optimization is triggered) can currently run. You may also notice that the Impala executors are not running.

Make sure that the /etc/resolv.conf file on the ECS hosts contain a maximum of 2 search domains.

Known issues in Database Catalogs

This topic describes the Database Catalog known issues for Cloudera Data Warehouse Private Cloud.

Known issues identified in 1.5.4

No new known issues identified in 1.5.4.

Known issues identified in 1.5.1

DWX-15302: Upgrade button stays visible even after the upgrade completes

After you upgrade the Database Catalog, the Upgrade button remains visible on the CDW web interface instead of disappearing or getting disabled.

Refresh the page on your browser.

Known issues in Hive Virtual Warehouses

This topic describes the Hive Virtual Warehouse known issues for Cloudera Data Warehouse (CDW) Private Cloud.

Known issues identified in 1.5.4

DWX-18505: Resource template changes back to Reduced resources while creating a Virtual Warehouse

On the **Create Virtual Warehouse** modal, suppose you selected Default resources from the Resource Template drop-down menu. When you wait for a few seconds or click elsewhere on the modal, you see that the resource template changes back to "Reduced resources". However, when you create the Virtual Warehouse, the Virtual Warehouse is created with the resource template you initially selected. This is a CDW UI defect you encounter on environments that are activated using the low resource mode.

None.

DWX-18445: Virtual Warehouses get stuck in the deleting state

You may intermittently notice that a Virtual Warehouses gets stuck in the "deleting" state when you delete a Virtual Warehouse in CDW. This could be due to a failing diagnostic job that is in an "Init:0/1" state in the namespace.

None. The diagnostic job gets cleaned up in one hour, after which the Virtual Warehouse gets deleted successfully.

Known issues identified in 1.5.2

DWX-16989: Hive query running on Iceberg table fails randomly

Suppose you have disabled the auto-suspend option for a Hive Virtual Warehouse or if the Virtual Warehouse is under continuous load and hence it cannot be stopped by the auto-suspend option. In this situation, using the Iceberg table format may cause the following exceptions to appear in the query coordinator log along with the submitted queries that have failed:

```
org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.token.SecretMar
$InvalidToken): token (token for hive: HDFS_DELEGATION_TOKEN
 owner=hive/dwx-env-host-1.cdp.local@EXAMPLE.CLOUDERA.COM,
 renewer=hive, realUser=, issueDate=1709813340891,
 maxDate=1710418140891, sequenceNumber=19784486, masterKeyId=52)
 is expired, current time: 2024-03-08 04:09:32,835-0800 expected
 renewal time: 2024-03-08 04:09:00,891-0800
```

```
org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.token.SecretMar
$InvalidToken): token (token for hive: HDFS_DELEGATION_TOKEN
 owner=hive/dwx-env-host-1.cdp.local@EXAMPLE.CLOUDERA.COM,
 renewer=hive, realUser=, issueDate=1699855596578,
 maxDate=1700460396578, sequenceNumber=16863242, masterKeyId=39)
 can't be found in cache
```

```
org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.token.SecretMar
$InvalidToken): token (OzoneToken owner=hive/dwx-env-
ewxf6g-env.cdp.local@ROOT.EXAMPLE.SITE, renewer=hive,
 realUser=, issueDate=2024-03-19T21:49:31.033Z,
 maxDate=2024-03-19T21:50:31.033Z, sequenceNumber=72,
```

```
masterKeyId=1, strToSign=null, signature=null,
awsAccessKeyId=null, omServiceId=ozone1710521984,
omCertSerialId=11) is expired, current time: 2024-03-19
21:51:34,293+0000 expected renewal time: 2024-03-19
21:51:31,033+0000
```

```
org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.token.SecretMan
$InvalidToken): token (OzoneToken owner=hive/dwx-env-
azt1gg-env.cdp.local@ROOT.EXAMPLE.SITE, renewer=hive,
realUser=, issueDate=2024-04-09T16:04:12.889Z,
maxDate=2024-04-09T17:04:12.889Z, sequenceNumber=29,
masterKeyId=1, strToSign=null, signature=null,
awsAccessKeyId=null, omServiceId=ozone1711550158,
omCertSerialId=2597525731772327) can't be found in cache
```

This happens because the HDFS delegation tokens are not renewed when using the Iceberg table
format. After the existing HDFS delegation tokens expire, Hive query coordinator (TEZ App
Master) cannot access the tables on the file system during the query planning phase. The problem
is independent of the file system--Ozone FS or Hadoop FS. The error only occurs after the HDFS
delegation tokens have expired. By default, the delegation tokens expire in one day. However, you
can modify the expiration time on the CDP Base cluster.

The problem does not occur if the query coordinator pods in the Hive Virtual Warehouse are
stopped manually or by using the auto-suspend functionality within the token expiration period.

Apply this workaround only if you cannot suspend the Hive Virtual Warehouse.

1. Log in to the Data Warehouse service as DWAdmin.
2. Go to the Virtual Warehouses tab and click ⋮ Edit Configurations Query Coordinator .
3. Select env from the Configuration files drop-down menu.
4. Add the following value against the JVM_OPTS property:

```
-Diceberg.scan.plan-in-worker-pool=false
```

5. Click Apply Changes.

Known issues identified in 1.5.1

DWX-15480: Hive queries fail with FILE_NOT_FOUND error

ACID directory cache may become outdated in Tez AMs in case of ACID tables that change often,
possibly leading to different errors with the same root cause: "split generation works from cache
pointing to non-existing files". And you may see the following error in the diagnostic bundles and
query logs: FILE_NOT_FOUND: Unable to get file status.

Disable the cache by setting the value of the hive.txn.acid.dir.cache.duration property to -1 by going
to Virtual Warehouse ⋮ Edit CONFIGURATIONS Hue Configuration files hive-site from the
CDW web interface.

DWX-15287: Drop database query for Hive fails with Invalid ACL Exception

You may see the following error in a Hue or beeline session when running DROP DATABASE,
DROP TABLE, or ALTER TABLE DROP PARTITION operations on a Hive Virtual Warehouse
that is in Stopped state: "org.apache.zookeeper.KeeperException$InvalidACLException:
KeeperErrorCode = InvalidACL for /llap-sasl/user-hive".

The exception is caused because the Hive Virtual Warehouse tries to evict the cache in the LLAP
executors, but the compute pods of the stopped warehouse are no longer running.

> **Note:** The database or table is deleted despite the exception, only the LLAP
> executors do not flush their database or table related buffers, because these executors
> are not running.

Start the Virtual Warehouse before you run the DROP DATABASE, DROP TABLE, or ALTER
TABLE DROP PARTITION operations.

Alternatively, you can add the hive.llap.io.proactive.eviction.enabled=false setting in the hive-sit
e.xml file. This method may result in some performance degradation, because LLAP no longer
discards the dropped database/table or temp table related buffers.

1. Log in to CDW as DWAdmin.

2. Click ⋮ Edit CONFIGURATIONS Hiveserver2 on the Virtual Warehouse tile and select
hive-site from the Configuration files drop-down menu.

3. Click ＋ and add the following line:

```
hive.llap.io.proactive.eviction.enabled=false
```

4. Click Apply Changes.

Wait for the Virtual Warehouse to refresh and return to Running or Stopped state.

Known issues in Impala Virtual Warehouses

This topic describes the Impala Virtual Warehouse known issues for Cloudera Data Warehouse Private Cloud.

Known issues identified in 1.5.4

DWX-18505: Resource template changes back to Reduced resources while creating a Virtual Warehouse

On the **Create Virtual Warehouse** modal, suppose you selected Default resources from the
Resource Template drop-down menu. When you wait for a few seconds or click elsewhere on the
modal, you see that the resource template changes back to "Reduced resources". However, when
you create the Virtual Warehouse, the Virtual Warehouse is created with the resource template you
initially selected. This is a CDW UI defect you encounter on environments that are activated using
the low resource mode.

None.

DWX-18445: Virtual Warehouses get stuck in the deleting state

You may intermittently notice that a Virtual Warehouses gets stuck in the "deleting" state when
you delete a Virtual Warehouse in CDW. This could be due to a failing diagnostic job that is in an
"Init:0/1" state in the namespace.

None. The diagnostic job gets cleaned up in one hour, after which the Virtual Warehouse gets
deleted successfully.

Known issues identified in 1.5.1

DWX-14292: Impala executors and coordinator pods get stuck in pending state

If you have enabled low resource mode while activating an environment in CDW and also enabled
the option to customize pod sizes for Impala from the **Advanced Settings** page, then you may
notice that the executors and coordinator pods do not request the set amount of memory and CPU.
This happens because the low resource mode overrides the default pod configuration.

Low resource mode and custom pod configurations for Impala cannot be used at the same time. To
resolve this issue, do not select the Low resource mode option while activating an environment.
Instead, create custom pod configurations with lower resources for Impala. You can continue to use
lesser resources as defined in the custom pod configuration for Hive.

VIZ-2046: Trusted impersonation does not work in CDW

When you try to create a connection to the Impala Virtual Warehouse from Cloudera Data
Visualization (CDV), CDV is unable to use Trusted Impersonation as a mode of authentication.
Therefore, you need to enter both the username and password for authentication.

None.

Known issues in Hue

This topic describes the Hue known issues for Cloudera Data Warehouse (CDW) Private Cloud.

Known issues identified in 1.5.4

CDPD-69394: Hue does not display logs while creating a table using the Importer

When you import a file in Hue using the Importer to create a table, the create table query is triggered and the operation succeeds. However, you may intermittently notice Hue not displaying the logs on the web interface.

None. This issue does not impact the underlying operation.

Known issues identified in 1.5.1

CDPD-54376: Clicking the home button on the File Browser page redirects to HDFS user directory

When you are previewing a file on any supported filesystem, such as S3, ABFS, or Ozone and you click on the Home button, you are redirected to the HDFS user home directory instead of the user home directory on the said filesystem.

None.

DWX-15090: Intermittently see the CSRF error in the Hue Job Browser

You may intermittently see the "403 - CSRF" error on the Hue web interface as well as in the Hue logs after running Hive queries from Hue.

Reload the page or start a new Hue session.

DWX-14911: Export operation failing in a partitioned Hive table

You may see the following error in the Hive server logs when running the EXPORT TABLE query in Hive: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.ExportTask. Distcp operation failed. This issue occurs when the number of files and the size of files exceeds the limits specified in the hive.exec.copyfile.maxsize and hive.exec.copyfile.maxnumfiles properties respectively, and the EXPORT TABLE query launches the distributed copy (`distcp`) job to copy files of HDFS.

Increase the limit of following properties in HiveServer2:

- hive.exec.copyfile.maxsize to maximum size of files in a directory in bytes
- hive.exec.copyfile.maxnumfiles to maximum number of files in a directory

OPSAPS-66903: Ozone HTTPFS address configuration contains placeholders instead of the real values

The Hue File Browser does not automatically support browsing the Ozone filesystem because the HTTPFS role configuration for the Ozone service contains placeholder text instead of real values. You can view this configuration by going to Cloudera Manager Ozone service Instances HTTPFS Gateway Processes and clicking on the ozone-conf/httpfs-site.xml configuration file.

You must manually configure Hue to access Ozone using the File Browser by following the instructions listed in Enabling browsing Ozone from Hue on CDW Private Cloud.

Known issues identified in 1.5.0

DWX-12616: Hue limitation in CDW Private Cloud

Following are the known limitations in Hue in CDW Private Cloud 1.5.0:

- Hue Importer does not support importing files more than 200 KB in size
- Hue File Browser does not work if the HDFS service on the base cluster is configured for high availability

None.

DWX-13865: Hue File Browser does not work with HDFS HA

Hue File Browser is not accessible or displays a 403 error when you click on File Browser from the left assist panel in CDW Private Cloud if HDFS is configured for High Availability on the base cluster. Currently, Hue in CDW cannot obtain the hostname and the port from the HttpFS service on the base cluster. This is a known limitation.

You must manually build and specify the WebHDFS URL for Hue in CDW to connect to the HttpFS service on the base cluster.

1. Log in to Cloudera Manager as an Administrator.
2. Go to Clusters Hive Configuration and note the value present in the Kerberos Principal field.

   This is the Hive service account name.
3. Go to Clusters HDFS HttpFS Advanced Configuration Snippet (Safety Valve) for httpfs-site.xml and click **+** to add the following lines:

   Name: httpfs.proxyuser.hive.hosts, Value: *

   Name: httpfs.proxyuser.hive.groups, Value: *

   Replace hive with the actual Hive service account name.
4. Click Save Changes and restart the HDFS service or the HttpFS role.
5. Go to the Instances tab and note the hostname of the HttpFS role.
6. Go to the Configuration tab and note the port for the hdfs.httpfs.http.port configuration from the RESTPort field.

   The default value of the hdfs.httpfs.http.port configuration is 14000.
7. Use the hostname and the port to construct the WebHDFS URL as follows:

   ```
   https://[***HOSTNAME***]:[***PORT***]/webhdfs/v1
   ```

8. Log in to the Data Warehouse service as a DWAdmin.
9. Go to the Virtual Warehouse from which you want to connect Hue to the base cluster's HDFS service and click ⋮ Edit .
10. Go to CONFIGURATIONS Hue , select hue-safety-valve from the Configuration files drop-down list and add the following lines:

    ```
    [hadoop]
    [[hdfs_clusters]]
    [[[default]]]
    webhdfs_url=https://[***HOSTNAME***]:[***PORT***]/webhdfs/v1
    ```

    Specify the WebHDFS URL that you constructed earlier.
11. Click Apply changes.

    Update the hue-safety-valve configuration for any Hive or Impala Virtual Warehouses from which you want to connect to the base cluster HDFS.

Known issues identified before 1.4.1
DWX-9373: Unable to log into Hue as a local administrator

If you have logged into the CDP Management Console as a local administrator, then you may not be able to log into Hue and you may see the following error: "User is not authorized".

To access Hue, you must add your username to your organization's LDAP tree or log into the CDP Management Console using your LDAP credentials and then access Hue.

Known issues in Unified Analytics

This topic describes the Unified Analytics known issues for Cloudera Data Warehouse Private Cloud.

Known issues identified in 1.5.4

DWX-18475: Changing the resource template does not update the query executor and coordinator pods

If you have enabled ETL Isolation on an Unified Analytics Virtual Warehouse, changing the resoure template does not update the query executor and coordinator pods. They retain the initial values from the resource template that was used when you created the Virtual Warehouse.

Create a new Virtual Warehouse with the desired resource template.

DWX-18143: Unified Analytics-specific pods do not get updated after changing the resource template

When you change the resource template of an Impala Virtual Warehouse with the Unified Analytics option enabled from the **Virtual Warehouses Details** page, CDW only updates the resources for Impala coordinator, statestore, catalogd pods, and so on. It does not update the resources for the HiveServer2 (HS2), Hive query executor, Hive coordinator, and standalone query executor pods.

None.

Iceberg-related known issues in CDW Private Cloud

This topic describes the Iceberg-related known issues in Cloudera Data Warehouse (CDW) Private Cloud.

Known issues identified in 1.5.4

No new known issues identified in 1.5.4.

Known issues identified in 1.5.2

CDPD-59413: Unable to view Iceberg table metadata in Atlas

You may see the following exception in the Atlas application logs when you create an Iceberg table from the CDW data service associated with a CDP Private Cloud Base 7.1.8 or 7.1.7 SP2 cluster: Type ENTITY with name iceberg_table does not exist. This happens because the Atlas server on CDP Private Cloud Base 7.1.8 and 7.1.7 SP2 does not contain the necessary, compatible functionality to support Iceberg tables. This neither affects creating, querying, or modifying of Iceberg tables using CDW nor does it affect creating of policies in Ranger.

On CDP Private Cloud Base 7.1.9, Iceberg table entities are not created in Atlas. You can ignore the following error appearing in the Atlas application logs: ERROR - [NotificationHookConsumer thread-1:] ~ graph rollback due to exception (GraphTransactionInterceptor:200) org.apache.atlas.exception.AtlasBaseException: invalid relationshipDef: hive_table_storagedesc: end type 1: hive_storagedesc, end type 2: iceberg_table

If you are on CDP Private Cloud Base 7.1.7 SP2 or 7.1.8, then you can manually upload the Iceberg model file z1130-iceberg_table_model.json in to the /opt/cloudera/parcels/CDH/lib/atlas/models/1000-Hadoop directory as follows:

1. SSH into the Atlas server host as an Administrator.
2. Change directory to the following:

```
cd /opt/cloudera/parcels/CDH/lib/atlas/models/1000-Hadoop
```

3. Create a file called 1130-iceberg_table_model.json with the following content:

```
{
  "enumDefs": [],
  "structDefs": [],
  "classificationDefs": [],
```

```
  "entityDefs": [
    {
      "name": "iceberg_table",
      "superTypes": [
        "hive_table"
      ],
      "serviceType": "hive",
      "typeVersion": "1.0",
      "attributeDefs": [
        {
          "name": "partitionSpec",
          "typeName": "array<string>",
          "cardinality": "SET",
          "isIndexable": false,
          "isOptional": true,
          "isUnique": false
        }
      ]
    },
    {
      "name": "iceberg_column",
      "superTypes": [
        "hive_column"
      ],
      "serviceType": "hive",
      "typeVersion": "1.0"
    }
  ],
  "relationshipDefs": [
    {
      "name": "iceberg_table_columns",
      "serviceType": "hive",
      "typeVersion": "1.0",
      "relationshipCategory": "COMPOSITION",
      "relationshipLabel": "__iceberg_table.columns",
      "endDef1": {
        "type": "iceberg_table",
        "name": "columns",
        "isContainer": true,
        "cardinality": "SET",
        "isLegacyAttribute": true
      },
      "endDef2": {
        "type": "iceberg_column",
        "name": "table",
        "isContainer": false,
        "cardinality": "SINGLE",
        "isLegacyAttribute": true
      },
      "propagateTags": "NONE"
    }
  ]
}
```

4. Save the file and exit.

5. Restart the Atlas service using Cloudera Manager.

Fixed issues in Cloudera Data Warehouse Private Cloud

Review the issues fixed in this release of the Cloudera Data Warehouse (CDW) service.

DWX-16589: CDW does not display an error when Virtual Warehouses do not have sufficient quota to scale up

> When a Virtual Warehouse did not have enough quota to scale up, the autoscaling failed. But CDW did not display any error on the CDW web interface and you had to manually configure the autoscaling options according to the available quota. This issue has been resolved.

VIZ-2269: CDW Virtual Warehouse connection to Data Visualization error

> The "not authorized to delegate" error caused by impersonation configurations no longer occurs when creating or editing the Virtual Warehouse connection. For more information, see Data Visualization release notes.

DWX-16591: Concurrent merge and update Iceberg queries are failing

> Earlier, concurrent merge and update Iceberg queries failed with the following error in the Hive application logs: "Base metadata location hdfs://*[***LOCATION-A***]* is not same as the current table metadata location '*[***LOCATION-B***]*' for default.merge_insert_target_iceberg \rorg.apache.iceberg.exceptions.CommitFailedException". This issue has been fixed.

DWX-14292: Impala executors and coordinator pods get stuck in pending state

> Earlier, the low resource mode overrode the default pod configuration and therefore, the executors and coordinator pods could not request the set amount of memory and CPU. You had to activate the environment without selecting the Low resource mode and instead create custom pod configurations with lower resources for Impala. This issue has been resolved.

DWX-15100: Impala doesn't release the query profiles until the session is closed by the client

> The queries you submitted from Hue were stuck in the "Waiting to be closed" state and that the last query submitted was not displayed on the **Job Browse** page in Hue. This issue has been fixed.

Technical Service Bulletins

TSB 2024-745: Impala returns incorrect results for Iceberg V2 tables when optimized operator is being used in CDW

> For the latest update on this issue see the corresponding Knowledge Article: TSB 2024-745: Impala returns incorrect results for Iceberg V2 tables when optimized operator is being used in CDW.

TSB 2024-758: Truncate command on Iceberg V2 branches cause unintentional data deletion

> For the latest update on this issue see the corresponding Knowledge Article: TSB 2024-758: Truncate command on Iceberg V2 branches cause unintentional data deletion.

Deprecation notices in CDW Private Cloud

Review the features and functionalities that have been or will be removed or deprecated in this release of Cloudera Data Warehouse (CDW) Private Cloud.

The "Keep current image version" option has been removed

Earlier, when you rebuilt the Database Catalog or Virtual Warehouse, you had the option to retain the image version by selecting the Keep current image version option or upgrade to the latest version. This option has been removed from the CDW web interface. Now when you rebuild the Database Catalog or Virtual Warehouse, they always retain the existing image version.

Delegation user has been removed

A proxy user, called the delegation user which was required for impersonating authorization requests from Hue and Data Visualization to the Impala coordinator has been deprecated. During environment activation, you no longer need to specify a delegation username and password. CDW automatically sets the older delegation user to "impala" when you upgrade, rebuild, or refresh an Impala Virtual Warehouse or the Data Visualization instance.

In the existing Data Visualization connections, you must rename the proxy user (delegation user) to "impala" user. You can rename the delegation username manually or refresh, upgrade, or rebuild the Virtual Warehouse or the Data Visualization instance after upgrading to CDP Private Cloud Data Services 1.5.4.

The ability to create custom Database Catalogs has been removed

CDW no longer supports custom Database Catalogs. The Create multiple Database Catalogs option used for allowing you to create custom Database Catalogs has been removed from the **Advanced Configurations** page.

After you upgrade to CDP Private Cloud Data Services 1.5.4, the existing custom Database Catalogs remain until you deactivate the environment. You can continue to upgrade, refresh, and rebuild the existing custom Database Catalogs. However, you can no longer create new custom Database Catalogs.

Version information for Cloudera Data Warehouse Private Cloud components

Cloudera Data Warehouse (CDW) uses Hive, Impala, and Hue as its Runtime components and also provides integration with Cloudera Data Visualization. Review the version information of CDW Private Cloud 1.5.4 components.

Table 1: CDW Private Cloud version information

CDW component	Version
Hive	2024.0.17.3
Impala	2024.0.17.3
Hue	2024.0.17.3
Cloudera Data Visualization	7.2.2
CDW server	1.9.0
CDP CLI	0.9.115

Apache Iceberg version information

The following table shows the version of the Iceberg component in this release of CDW Private Cloud:

CDW server version	CDW Runtime version	Iceberg version
1.9.0	2024.0.17.3	1.3.0

Cumulative hotfixes: CDW Private Cloud

Review the cumulative hotfixes that have been shipped for Cloudera Data Warehouse Private Cloud 1.5.4.

CDW Private Cloud 1.5.4-CHF1

Review the features, fixes, and known issues in the Cloudera Data Warehouse (CDW) 1.5.4 Cumulative hotfix 1 release.

Fixed issues in CDW Private Cloud 1.5.4-CHF1

Review the issues fixed in this cumulative hotfix release of Cloudera Data Warehouse (CDW) Private Cloud.

CDPD-68683: A table with thousands of columns gets hanged in Hue Table Browser

When opening a table in the Table Browser in Hue, individual network requests were being made to fetch comments for each of the columns of the table. In the case where a table had a large amount of columns, this lead to a saturated network. This issue has been fixed by making only one call to fetch all the column comments in one go.

DWX-18505: Resource template changes back to Reduced resources while creating a Virtual Warehouse

On the **Create Virtual Warehouse** modal, when you selected Default resources from the Resource Template drop-down menu, and after a few seconds the resource template changed back to Reduced resources. This issue has been fixed.

DWX-18151: Hue backup and restore fails on OpenShift Container Platform (OCP)

Hue failed to connect to the Hue database, and the backup/restore operation failed because of incorrect permissions set for the pgpass file on mount due to the restricted Security Context Constraint on OCP. This issue has been fixed.

DWX-18508: Memory scaling issues for Impala Statestored

Earlier, Impala Statestored pods did not scale in a Virtual Warehouse created with a t-shirt group of 20 executors, and the Impala Virtual Warehouse ran into out of memory issues. This issue has been fixed. Now each Impala Statestored can scale their memory to 2 GB.

CDPD-70188: Hive-Iceberg queries fail with the IllegalStateException Invalid file error

Hive-Iceberg queries failed with the following error:

```
IllegalStateException Invalid file: file length 0 is less tha
 minimal length of the footer tail 12
```

This issue has been fixed.