

# Enabling Multi-tenancy in Cloudera Data Warehouse (Preview)

Date published: 2021-06-02

Date modified: 2021-06-02

## Legal Notice

© Cloudera Inc. 2021. 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.

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.

*This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided 'as is' without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.*

## Contents

<b>Overview</b>	<b>4</b>
<b>Creating multiple default Database Catalogs</b>	<b>5</b>

*This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided 'as is' without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.*

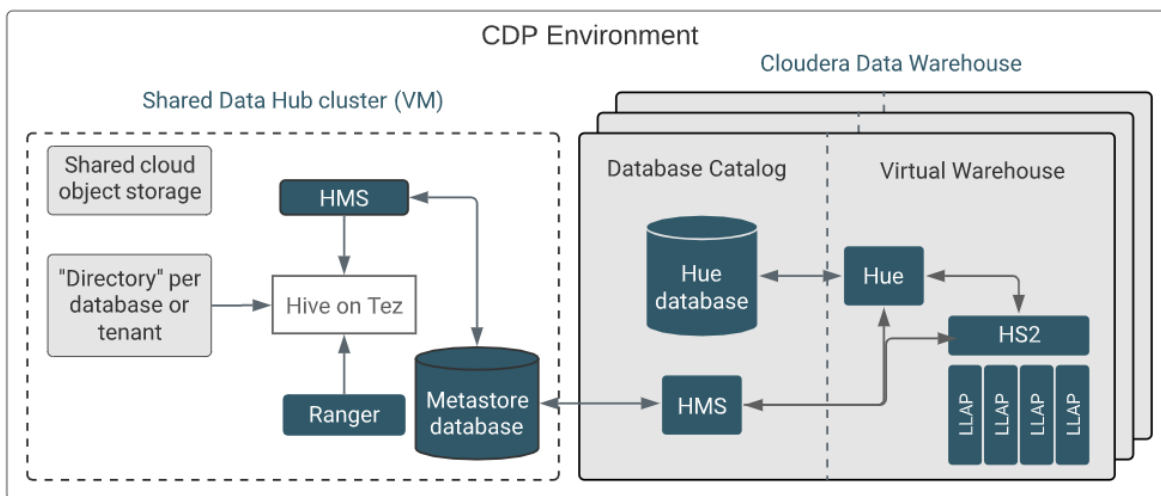
# Overview of multi-tenancy in Cloudera Data Warehouse

Some organizations use Cloudera Data Warehouse (CDW) to create and host proprietary data warehouses and data marts and provide business analytics service to their customers (or tenants) on their data using Hue. In this business model, the data provider must isolate the data sets and queries, job history, and other artifacts from its tenants for security, privacy, and resource isolation. CDW enables you to deploy a secure, multi-tenant environment for your customers by allowing you to create multiple default Database Catalogs within CDW. As a result, you can provide different Hue instances to each tenant.

Database Catalogs are Hive MetaStore (HMS) instances, and include references to the cloud storage where the data lives. An environment can have multiple Database Catalogs. When you activate an environment from the Data Warehouse UI, a default Database Catalog is created (format: environment\_name-default). This is the same HMS instance used by your CDP environment. You can add additional default Database Catalogs if you want a standalone data warehouse without any data from the tables that are in the environment.

The default Database Catalog shares the HMS database with HMS in the Data Hub cluster. This enables you to access any objects or data sets created in the Data Mart or the Data Engineering clusters from CDW Virtual Warehouses and vice versa.

The Hue database contains saved queries and query history and they are stored in the Database Catalog. By creating multiple default Database Catalogs, each tenant's query history, job details, saved queries, and so on, are isolated from the others. The following architecture diagram explains this concept:



If your data is stored on cloud object storage such as S3 or ABFS, then the access to the shared data is controlled on the database level for each tenant through a common Ranger

*This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided 'as is' without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.*

repository. This allows the tenants to have dedicated databases but you can also share data with them if required.

## Creating a new default Database Catalog

By creating multiple default Database Catalogs, each tenant gets a dedicated default Database Catalog and a Hive instance that serves the same metadata from the Shared Data Lake (SDX) but provides a tenant-specific Hue database. Each tenant also gets their own, fully functional Virtual Warehouse with dedicated compute resources.

To create a new default Database Catalog:

1. Log in to the CDP web interface and navigate to the Data Warehouse service.
2. Go to the Database Catalog page and click **ADD NEW**.
3. On the **New Database Catalog** dialog, specify the Database Catalog name, select an environment, and select **SDX** from the **Datalake** dropdown.
4. Click **CREATE** to create the new default Database Catalog.