Cloudera Data Warehouse Private Cloud 1.5.0

Cloudera Data Warehouse Private Cloud Environments

Date published: 2020-08-17
Date modified: 2023-01-25

CLOUDERA

Legal Notice

Contents

Overview of Private Cloud environments

An environment is a logical entity that represents the association of your Private Cloud user account with compute resources. Using the compute resources of the environment, you can provision and manage Cloudera Data Warehouse (CDW), Cloudera Data Engineering (CDE), or Cloudera Machine Learning (CML) workloads.

For a CDP Private Cloud deployment, you can configure the environment on an OpenShift or an Embedded Container Service (ECS) cluster. Deploying CDP Private Cloud Data Services (CDW, CDE, and CML) on OpenShift requires you to deploy and manage the Kubernetes infrastructure. ECS service creates and manages an embedded Kubernetes infrastructure for use with the CDP Private Cloud Experiences. To deploy CDP Experiences on ECS clusters, you only need to provide hosts on which you want to install the service and Cloudera Manager sets up an ECS cluster using an embedded Kubernetes service and provides a framework to manage and monitor the cluster.

The following diagram shows the components of CDP Private Cloud:
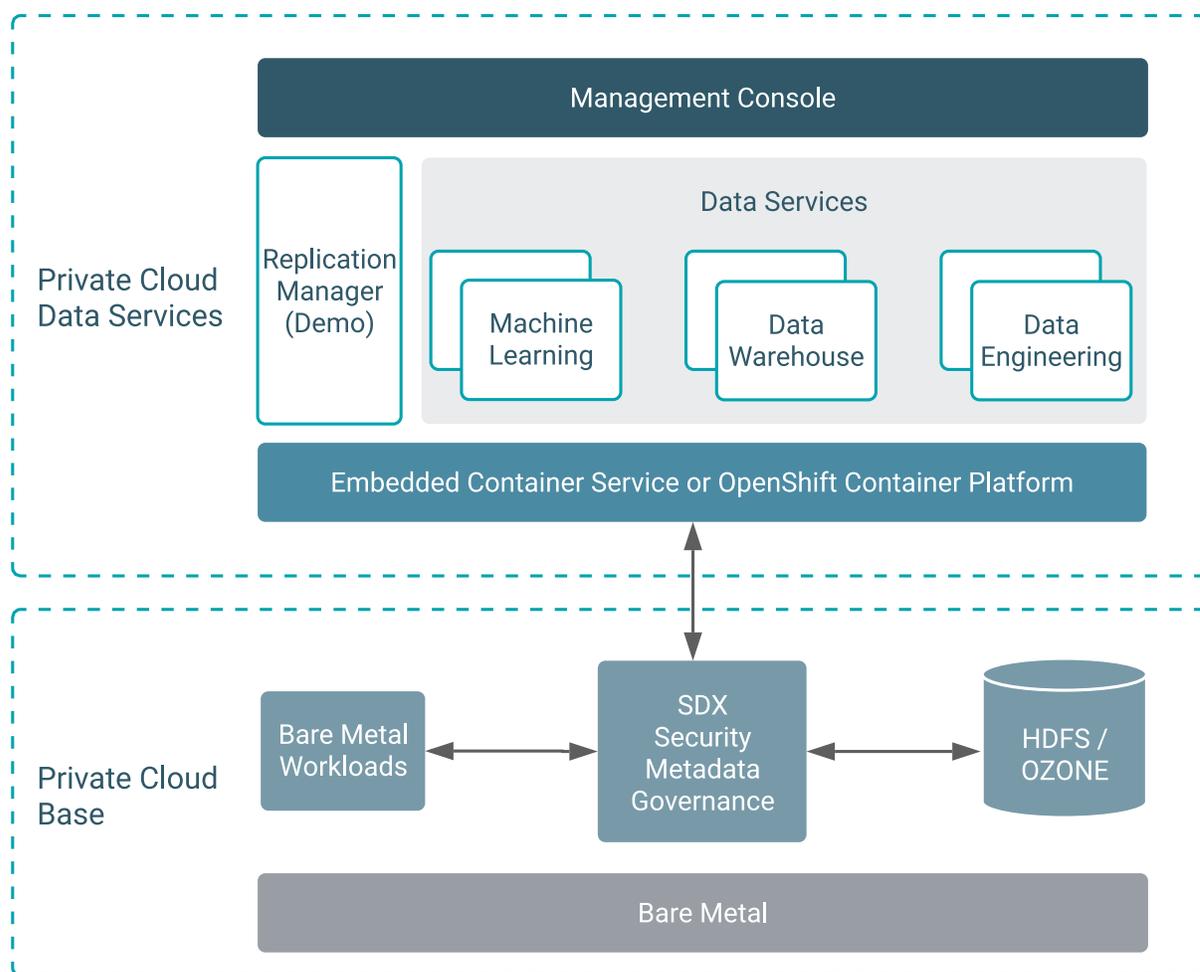
For more details about environment requirements and how to register an environment so CDP can access Kubernetes clusters on your OpenShift or ECS deployment, see CDP Private Cloud Environments.

Activating OpenShift environments

This topic describes how to activate an environment to use for Cloudera Data Warehouse (CDW) Private Cloud on Red Hat OpenShift Container Platform (OCP).

About this task

Before you can create a Database Catalog to use with a Virtual Warehouse, you must activate a CDP environment. Activating an environment causes CDP to connect to the Kubernetes cluster, which provides the computing resources for the Database Catalog. In addition, activating an environment enables the Cloudera Data Warehouse (CDW) service to use the existing data lake that was set up for the environment, including all data, metadata, and security.

Before you begin

- Determine which environment that uses a particular data lake is the environment you want to activate for use with a Database Catalog and Virtual Warehouse.
- For local caching, ensure that an administrator uses the Local Storage Operator to create a local file system on an SSD/NVMe for each OpenShift worker node and then mounts it to a known location on the worker node. Make sure that this local caching location allows temporary data to be stored in a way that supports performance. You need to specify the Storage Class Name from the Local Storage Operator when you activate the environment for the CDW service in Step 4 below. For more information about creating a local file system on OpenShift worker nodes using the Local Storage Operator, see Persistent storage using local volumes in the OpenShift documentation.

  Note: If you have more than one OCP clusters managed using different instances of Cloudera Manager, but using the same AD server and using the same environment name, then go to the Advanced Configuration Advanced Settings page and disable the Use deterministic namespace names option before activating the environment in CDW.

Procedure

1. Log in to Data Warehouse service as DWAdmin.
2. Expand the Environments column by clicking More… and locate the Environment that you want to activate.

3. Click the activation icon.

The **Activation Settings** dialog box is displayed.

Activation Settings   ✕

Do you want to activate the environment "dummy-openshift-env"?

Storage Class Name from Local Storage Operator *

> *Enter Storage Class Name*

Not a valid name

Security Context Constraint Name (optional)

> *Enter Security Context Constraint Name*

Delegation Username* ⓘ   Delegation Password*

> *Delegation Username*   *Delegation Password*

Pre-created database names for default database catalog* ⓘ

> *DAS*   *Hue*

◯ Enable Low Resource Mode

CANCEL   ACTIVATE

4. Specify the Storage Class Name from Local Storage Operator:

This is the Storage Class Name you specified when you created the local file system for caching as described in the above section Before you begin. It is the location where temporary data is stored.

> ⚠️ **Important:** Be sure to specify the correct Storage Class Name when activating an environment. If an incorrect Storage Class Name is specified, the environment might activate successfully, but Virtual Warehouses that use the environment do not start.

Optionally, you can specify the Security Context Constraint Name.

5. Specify Delegation Username and Delegation Password to impersonate authorization requests from Hue to the Impala engine.

> **Note:**
> - The delegation user and the LDAP Bind user configured on the **Administration** page of the Management Console are not necessarily the same user.
> - The special characters used in the LDAP Bind user password are not exactly the same as the ones that can be used in the delegation user password, because only the following characters are supported to be used in the LDAP Bind user password: ! # $ % ( ) * + , - . / : ; = ? @ [ ] ^ _ ` { | } ~.
> - The following special characters are not supported to be used in the name of the delegation user or in the Distinguished Name of the LDAP Bind user: < > & ' ".

6. Enable low resource mode to deploy CDW on minimum hardware.

> **Note:** Cloudera recommends that you use the Low Resource Mode option for proof of concept (POC) purposes only. This feature is not recommended for production deployment.
>
> Complex queries and multiple queries on HS2 may fail due to limited memory configurations for HMS and HS2 in the low resource mode.

7. Specify Delegation Username and Delegation Password to impersonate authorization requests from Hue to the Impala engine.

8. If you are using an external database on your base cluster and want to use a default Database Catalog, then you must specify custom database name for DAS and Hue in the Pre-created database names for default database catalog field.

9. Click ACTIVATE.

Related Information

[Adding a new Virtual Warehouse](#)

Activating Embedded Container Service environments

This topic describes how to activate an environment to use for Cloudera Data Warehouse (CDW) Private Cloud on Embedded Container Service (ECS).

About this task

Before you can create a Database Catalog to use with a Virtual Warehouse, you must activate a CDP environment. Activating an environment causes CDP to connect to the Kubernetes cluster, which provides the computing resources for the Database Catalog. In addition, activating an environment enables the Cloudera Data Warehouse (CDW) service to use the existing data lake that was set up for the environment, including all data, metadata, and security.

Before you begin

- Determine which environment that uses a particular data lake is the environment you want to activate for use with a Database Catalog and Virtual Warehouse.
- In ECS environments, the Storage Class Name is automatically obtained from Cloudera Manager.

> **Note:** A "default" environment is created by the Control Plane when you add a Private Cloud cluster. If you have more than one ECS clusters managed using different instances of Cloudera manager, but using the same Active Directory (AD) server and using the same "default" environment, then go to the Advanced Configuration Advanced Settings page and disable the Use deterministic namespace names option before activating the environment in CDW.

Procedure

1. Log in to Data Warehouse service as DWAdmin.

2. Expand the Environments column by clicking More… and locate the Environment that you want to activate.

3. Click the activation icon.

   The **Activation Settings** dialog box is displayed.

4. Specify Delegation Username and Delegation Password to impersonate authorization requests from Hue to the Impala engine.

   > **Note:**
   > - The delegation user and the LDAP Bind user configured on the **Administration** page of the Management Console are not necessarily the same user.
   > - The special characters used in the LDAP Bind user password are not exactly the same as the ones that can be used in the delegation user password, because only the following characters are supported to be used in the LDAP Bind user password: ! # $ % ( ) * + , - . / : ; = ? @ [ ] ^ _ ` { | } ~.
   > - The following special characters are not supported to be used in the name of the delegation user or in the Distinguished Name of the LDAP Bind user: < > & ' ".

5. If you are using an external database on your base cluster and want to use a default Database Catalog, then you must specify custom database name for DAS and Hue in the Pre-created database names for default database catalog field.

6. Enable low resource mode to deploy CDW on minimum hardware.

> **Note:** Cloudera recommends that you use the Low Resource Mode option for proof of concept (POC) purposes only. This feature is not recommended for production deployment.
>
> Complex queries and multiple queries on HS2 may fail due to limited memory configurations for HMS and HS2 in the low resource mode.

7. Click ACTIVATE.

Related Information

Adding a new Virtual Warehouse

Authentication using Delegation Username and Password

You must specify a delegation username and password to impersonate authorization requests from Hue to the Impala engine during environment activation. The delegation user and password can authenticate users through an LDAP service account.

The ability to specify an LDAP delegation user also allows you to freely use special characters in your LDAP Bind DN, as CDW no longer has to inherit and process the delegation user from the LDAP Bind DN.

> **Note:**
> - The delegation user and the LDAP Bind user configured on the **Administration** page of the Management Console are not necessarily the same user.
> - The special characters used in the LDAP Bind user password are not exactly the same as the ones that can be used in the delegation user password, because only the following characters are supported to be used in the LDAP Bind user password: ! # $ % ( ) * + , - . / : ; = ? @ [ ] ^ _ ` { | } ~.
> - The following special characters are not supported to be used in the name of the delegation user or in the Distinguished Name of the LDAP Bind user: < > & ' ".

You can change the delegation username and password even after activating the environment.

The following image shows the CDW **Activation Settings** page containing the Delegation Username and Delegation Password fields:

Activation Settings   ✕

Do you want to activate the environment "dummy-openshift-env"?

Storage Class Name from Local Storage Operator *

    Enter Storage Class Name

Not a valid name

Security Context Constraint Name (optional)

    Enter Security Context Constraint Name

Delegation Username* ⓘ   Delegation Password*

    Delegation Username   Delegation Password

Pre-created database names for default database catalog* ⓘ

    DAS   Hue

○ Enable Low Resource Mode

                                                    CANCEL   ACTIVATE

Viewing environment details in Cloudera Data Warehouse Private Cloud

This topic describes how to view details of an environment created on an OpenShift cluster in Cloudera Data Warehouse (CDW) Private Cloud.

About this task

You can view environment details without leaving the CDW Private Cloud service UI. Accessing the Environment Details page also enables you to edit the description of the environment.

Before you begin

You must activate an environment before you can view its details. See "Activating environments," which is linked to in the "Related information" section at the bottom of this page.

Procedure

1. In the CDW Data Service, expand the Environments column by clicking the More… menu on the left side of the page.
2. In the Environments column, click the search icon and locate the environment you want to view.
3. When you locate the environment you want to view, click the edit icon in the upper right corner of the tile. This launches the Environment Details page.
4. In the Environment Details page, you can view information about the environment, like when it was created and last updated, or how many Database Catalogs and Virtual Warehouses use the environment.
5. If you changed the environment description, click Apply in the upper right corner of the page to save it.

Related Information

Activating OpenShift environments