

CDP Private Cloud Data Services Data Recovery

Date published: 2020-12-16

Date modified: 2023-01-25



Legal Notice

© Cloudera Inc. 2024. 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.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

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.

Contents

Data Recovery Service overview.....	4
Learn how backup and restore events work in Backup and Restore Manager.....	4
Access Backup and Restore Manager in Management Console.....	5
Backup Overview section on Management Console Dashboard.....	5
View Backup and Restore Manager.....	6
Backups tab in Backup and Restore Manager.....	6
Restores tab in Backup and Restore Manager.....	8
Creating backup of Control Plane and restoring it.....	10
Troubleshooting Backup and Restore Manager.....	11
CDP Control Plane UI or the Backup and Restore Manager becomes inaccessible after a failed restore event.....	11
Timeout error appears in Backup and Restore Manager.....	12
Stale configurations in Cloudera Manager after a restore event.....	12
Timeout error during backup of OCP clusters.....	13
Using DRS with CDW.....	13

Data Recovery Service overview

The Data Recovery Service (DRS) is a microservice in CDP Private Cloud Data Services that allows you to back up and restore Kubernetes namespaces and resources on both Embedded Container Service (ECS) and OpenShift Container Platform (OCP). You can backup and restore namespaces for Control Plane and other supported services. The "Backup and Restore Manager" in the CDP Private Cloud Data Services Management Console, and the CDP CLI options for the respective services leverage the DRS capabilities to run the backup and restore events (jobs).

Cloudera recommends that you create a backup of your Kubernetes namespace before a maintenance activity, before you upgrade, or in general, as a best practice.

Role required: *PowerUser*

By default, the data recovery service is located in the `[***CDP_INSTALLATION_NAMESPACE***]-drs` namespace. For example, if the CDP Private Cloud Data Services installation is located in the `cdp` namespace, the data recovery service namespace is automatically named `cdp-drs`. If you have multiple CDP Private Cloud Data Services installations (as in OCP), the data recovery service is named accordingly.

When you initiate the backup event in the Backup and Restore Manager for Control Plane, the data recovery service takes a backup of the following resources and data:

- Kubernetes resources associated with the `cdp` namespace and the embedded vault namespaces of the Control Plane in CDP Private Cloud Data Services. The resources include deployment-related information, stateful sets, secrets, and configmaps.
- Data used by the stateful pods, such as the data in the embedded database and Kubernetes persistent volume claim.

The Data Recovery Service requires CSI snapshots to back up and restore Kubernetes namespaces and resources. The CSI snapshots are enabled on ECS by default. You might require an additional license to enable CSI snapshots in Red Hat ODF storage on OCP.

Alternatively, you can use the CDP CLI options in [CDP CLI drscp](#) to backup and restore Control Plane.

Learn how backup and restore events work in DRS

The "Backup and Restore Manager" in the CDP Private Cloud Data Services Management Console and the CDP CLI options use Data Recovery Service (DRS) to backup and restore Kubernetes namespaces and resources on Embedded Container Service (ECS) and OpenShift Container Platform (OCP) for Control Plane and other supported services. You can take multiple backups and restore any backup as necessary.

Backup event

The backup event does not have any downtime impact and you can backup the Control Plane while it is running.

When you create a backup, the data recovery service:

1. initiates the backup event or job for the chosen backup entity. For example, the Control Plane in CDP Private Cloud Data Services,
2. assigns an ID called `backupCrn` to the backup event,

The `backupCRN` appears in the CRN column on the Backup and Restore Manager Backups tab. Click the **CRN** to view more details about the backup event on the Backup `[***name of backup***]` modal window.

3. creates a backup of the persistent volume claim (PVC) snapshots of the Control Plane namespaces and the backup event's PVC.



Tip: Data recovery service is a microservice in CDP Private Cloud Data Services that backs up and restores the Kubernetes namespaces and resources of supported services. Backup and Restore Manager leverages the data recovery service capabilities to backup and restore namespaces in Management Console.

Restore event

Best practices:

- Do not delete the `[**CDP_INSTALLATION_NAMESPACE**]-drs` namespace while the restore event is in progress.
- If you want to restore a backup from version 1.5.0 to version 1.5.1, you must perform the following steps:
 1. Create a backup of alert rules in 1.5.1 using the `kubectl get alertruletemplates -o yaml > [***alert backup file name***].yaml` command. This is because alert rules are not supported in 1.5.0 and they get deleted when you restore a 1.5.0 backup in version 1.5.1.
 2. Restore the backup from version 1.5.0.
 3. Restore the alert rules backup that you created in *Step 1* using the `kubectl apply -f [***alert backup file name***].yaml` command.

When you start the restore event, the data recovery service:

1. initiates the restore event based on the chosen backup,
2. assigns an ID called `restoreCrn` to the restore event,

The `restoreCRN` appears as `CRN` on the `Backup and Restore Manager Restores` tab. Click the `CRN` to view more details about the restore event.

3. deletes the existing resources and data,

During this stage of the restore event, the ECS restore vault is sealed and the POD is down which might appear as a failure in the control plane environment. After the restore event is complete, the vault and POD are auto-recovered and restored. Depending on the number of resources and data, this step might take a maximum of 10 minutes to complete.

4. restores the resources and data from the backup.

The restore event has a downtime impact because the pods and data are recreated.

Related Information

[CDP CLI drscp](#)

Access Backup and Restore Manager in Management Console

To backup and restore Kubernetes namespaces and resources on Embedded Container Service (ECS) and OpenShift Container Platform (OCP), you can use the Backup and Restore Manager in the CDP Private Cloud Data Services Management Console.

You can view the list of available backup entities in the "Backup Overview" section on the Dashboard page of the CDP Private Cloud Data Services Management Console. When you click "View Details" in this section, the "Backup and Restore Manager" appears. You can create backups, and view and perform actions on the available backup and restores entities in the "Backups" tab and "Restores" tab respectively.



Tip: Data recovery service is a microservice in CDP Private Cloud Data Services that backs up and restores the Kubernetes namespaces and resources of supported services. Backup and Restore Manager leverages the data recovery service capabilities to backup and restore namespaces in Management Console.

Backup Overview section on Management Console Dashboard

You can view all the available backup entities for all the supported services, total number of backups available for each entity, and the status of each backup event in the "Backup Overview" section on the "Dashboard" page in CDP Private Cloud Data Services Management Console.

The following columns appear in the `Dashboard Backup Overview` section:

Column name	Description
Backup Entity	Lists all the backup entities that are available for backup. For example, the CDP Control Plane: [***namespace***] is the Control Plane backup entity.
Total Backups	Total number of backups for the backup entity.
Colored dots	Each dot pertains to the current status of a backup event or job. Hover over a dot to view the backup event start timestamp and its current status. The yellow dot denotes NOT_STARTED or IN_PROGRESS backup event state, the green dot denotes COMPLETED event state, and the red dot denotes PARTIALLY_FAILED or FAILED event state.

When you click View Details, the **Backup and Restore Manager** appears.

The following sample image shows the **Backup Overview** section:



View Backup and Restore Manager

The Backup and Restore Manager appears after you click "View Details" in the "Backup Overview" section on the "Dashboard" page in CDP Private Cloud Data Services Management Console. The "Backups" tab lists all the backups and the "Restores" tab lists all the restore events.

Click New Backup on the **Backup and Restore Manager** page to initiate a backup event. The page shows the backup entity name and provides the following tabs:

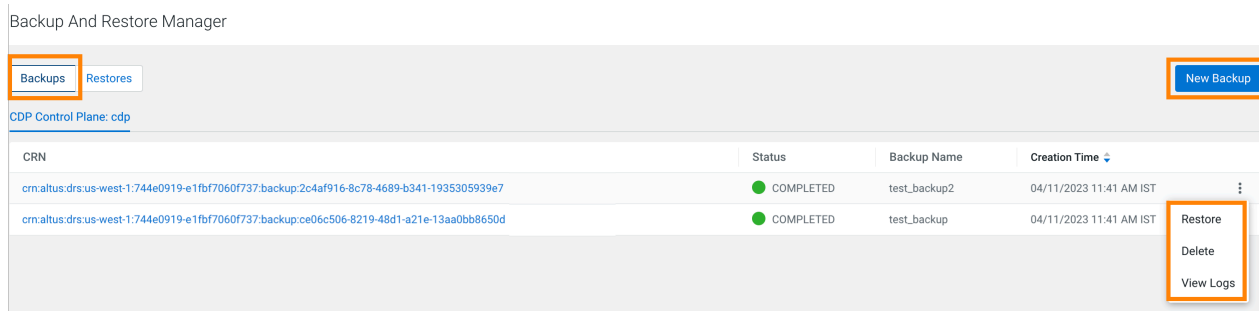
Backups tab in Backup and Restore Manager

The "Backups" tab in Backup and Restore Manager on the CDP Private Cloud Data Services Management Console lists all the available backups. You can create backups, or perform actions such as restore, delete, or view logs for each back up as necessary.

The following table lists the columns that appear on the **Backups** tab:

Column name	Description
CRN	Automatically assigned ID or backupCrn for the backup event. Customer Resource Number (CRN) is the <i>Cloudera</i> -specific identifier provided for the event/job. Click the CRN to view more details about the event on the Backup [***name of backup***] modal window.
Status	Current backup event status. The event states include NOT_STARTED, IN_PROGRESS, COMPLETED, PARTIALLY_FAILED, and FAILED.
Backup Name	Unique name given to the backup event while initiating the backup event.
Creation Time	Timestamp when the backup event was initiated.

The following sample image shows the **Backups** tab on the **Backup and Restore Manager** page:



You can perform the following actions on each successful backup event:

- Restore the backup.
- Delete the backup. This deletes the backup permanently.
- View Logs opens the Backup [***name of backup***] modal window.

On the Backup [***name of backup***] modal window, you can choose to Restore the backup, Delete the backup, or click Cancel to close the window. The window also shows the following tabs:

Tab	Description
Details	<ul style="list-style-type: none"> • CRN of the backup event • Creation Time and date • Completed or Updated Time and date • Current Status of the backup event • Backup Phase the event is running in (such as in-progress or finished) • Backup Name that was assigned to the backup event during creation • The Included Namespaces in the backup event
Logs	Provides the log details about the backup event.

The following sample image shows the Backup [***name of backup***] modal window:

Backup test_backup

[Details](#)[Logs](#)**CRN**

crn:altus:drs:us-west-1:744e0919-e1fbf7060f737:backup:ce06c506-8219-48d1-a21e-13aa0bb8650d

Creation Time

04/11/2023 11:41 AM IST

Updated Time

04/11/2023 11:42 AM IST

Status● COMPLETED**Backup Phase**

FINISHED

Backup Name

test_backup

Included Namespaces

vault-system, cdp

Restore

Delete

Cancel

Restores tab in Backup and Restore Manager

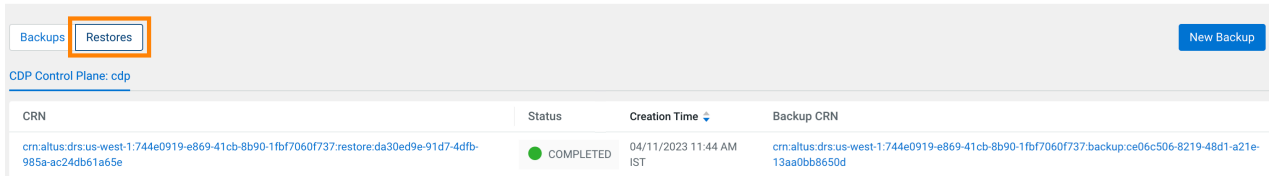
The "Restores" tab in Backup and Restore Manager on the CDP Private Cloud Data Services Management Console lists all the available restore events.

The following table lists the columns that appear on the **Restores** tab:

Column name	Description
CRN	Automatically assigned ID or restoreCrn for the restore event. When you click the CRN, the Restore Details modal window appears.
Status	Current restore event status as COMPLETED or FAILED.
Creation Time	Timestamp when the restore event is initiated.
Backup CRN	CRN of the backup event that is being restored. When you click the Backup CRN , the Backup [***name of backup***] modal window appears.

The following sample image shows the **Restores** tab on the **Backup and Restore Manager** page:

Backup And Restore Manager



When you click the CRN of a restore event, the following tabs appear on the **Restore Details** modal window:

Column name	Description
Details	<ul style="list-style-type: none"> • CRN of the restore event • Creation Time and date of the restore event • Completed or Updated Time and date of the restore event • Current Status of the restore event • The Restore Phase the event is running in (such as in-progress, pending, failed, or finished) • Associated Backup CRN of the backup event that was restored • The Included Namespaces in the restore event • Warnings or Errors. <p>When a warning appears, you can continue to use the backup or restore event. However, it is advisable to scrutinize the warning to avoid any potential issues. Errors appear if the restore event has failed.</p>
Logs	Provides the log details about the event.

The following sample image shows the Restore Details modal window:

Restore Details



[Details](#)

[Logs](#)

CRN

crn:altus:drs:us-west-1:017a9c10-d8e7-472f-88ed-158946a2fe84:restore:a6668773-defb-4f45-b65d-c2160055205a

Creation Time

04/27/2023 5:40 PM IST

Updated Time

04/27/2023 5:45 PM IST

Status

● COMPLETED

Restore Phase

FINISHED

Associated Backup CRN

crn:altus:drs:us-west-1:017a9c10-d8e7-472f-88ed-158946a2fe84:backup:a9bde036-5071-4f8f-afd7-141edffb7f9f

Included Namespaces

vault-system, cdp

Creating backup of Control Plane and restoring it

The Backup and Restore Manager in the CDP Private Cloud Data Services Management Console helps you to backup and restore Kubernetes namespaces and resources on Embedded Container Service (ECS) and OpenShift Container Platform (OCP). You can also restore and delete the backups.

Before you begin

Ensure that the following prerequisites are complete:

- You must have the *PowerUser* role.
- For OCP, ensure that a *VolumeSnapshotClass* is installed with a CSI driver that matches the CSI driver for the storage class used.



Tip: Data recovery service is a microservice in CDP Private Cloud Data Services that backs up and restores the Kubernetes namespaces and resources of supported services. Backup and Restore Manager leverages the data recovery service capabilities to backup and restore namespaces in Management Console.

Alternatively, you can use the data recovery service CDP CLI options for Control Plane to create, restore, or delete backups. For more information, see [CDP CLI drscp](#).

About this task

The following steps show how to create a backup of the Kubernetes namespaces and resources in the Control Plane and then restore it.

Procedure

1. Go to the [CDP Private Cloud Data Services Management Console Dashboard](#) page.
2. To create a backup, perform the following steps:
 - a) Click **Create Backup** in the **Backup Overview** section to create the first backup.
For subsequent backups, click **View Details Backup and Restore Manager New Backup** to create a backup.
 - b) Enter a unique **Backup Name** and choose the **Backup Entity** that you want to back up in the **Create Backup** modal window.
 - c) Click **Create**.
This initiates the backup event and generates a backupCRN. The backupCRN appears as a CRN on the **Backups** tab in the **Backup and Restore Manager** that you can click to view the backup event details.
3. To restore a backup, perform the following steps:
 - a) Go to the **Backup and Restore Manager Backups** tab.
 - b) Click **Actions Restore**, and then click **OK** in the **Restore** modal window to acknowledge that you want to restore the backup.
Alternatively, click the CRN of the required backup; click **Restore** on the **Backup [***name of backup***]** modal window, and then click **OK** to acknowledge that you want to restore the backup.
4. To delete a backup, perform the following steps:
 - a) Go to the **Backup and Restore Manager Backups** tab.
 - b) Click **Actions Delete**, and then click **OK** in the **Delete** modal window to acknowledge that you want to delete the backup.
Alternatively, click the CRN of the required backup; click **Delete** on the **Backup [***name of backup***]** modal window; and then click **OK** to acknowledge that you want to delete the backup.

Troubleshooting Backup and Restore Manager

The troubleshooting scenarios in this topic help you to troubleshoot issues that might appear for DRS in the Control Plane. The “Backup and Restore Manager” in CDP Private Cloud Data Services Management Console leverages the data recovery service capabilities to backup and restore Kubernetes namespaces and resources.

CDP Control Plane UI or the Backup and Restore Manager becomes inaccessible after a failed restore event

Condition

What to do if the CDP Control Plane UI does not come up or the Backup and Restore Manager (or drscp options) becomes inaccessible after a failed restore event?

Cause

Sometimes, some configurations take more time to restore. For example, in a shared cluster (OCP) that is heavily loaded, the restore event might surpass the set timeout limit. In this scenario, you can either wait or rerun the restore event again.



Tip: Run the restore event for such scenarios during non-peak hours.

Solution

You can perform one of the following steps after a failed restore event:

- Wait for a minimum of 15 minutes. This might resolve the issue automatically if the issue was caused due to timeout. You can verify this in the logs.
- Run restore again. This might resolve the issue if it was temporary such as, restore event during cluster maintenance.

If the Control Plane is not restored successfully even after you follow the steps, contact Cloudera Support for further assistance.

Timeout error appears in Backup and Restore Manager

Condition

What to do if a timeout error appears in the Backup and Restore Manager (or drscp options) during a restore event?

Solution

When the restore event crosses the time set in the `POD_CREATION_TIMEOUT` environment property of the `cdp-release-thunderhead-drsprovider` deployment in the `drs` namespace, a timeout error appears. By default, the property is set to 900 seconds. In this scenario, you must manually verify whether the pods are up or not.

Stale configurations in Cloudera Manager after a restore event

Condition

Why are stale configurations in Cloudera Manager found after a restore event?

Cause

This scenario appears when you take a backup of the CDP Private Cloud Data Services Control Plane, upgrade Data Services, and then perform a restore. During the upgrade process, new parcels are activated and configurations in Cloudera Manager might have changed.

Solution

It is recommended that you restart Cloudera Manager after the upgrade process is complete and then initiate the restore event.

Timeout error during backup of OCP clusters

Condition

What to do when the “The execution of the sync command has timed out” error appears during a backup event for OCP clusters?

Cause

This scenario is observed when the cluster is heavily used and the backup event is initiated during peak hours.

Solution

You can restart the nodes, this causes the disk to unmount and forces the operating system to write any data in its cache to the disk. After the restart is complete, initiate another backup. If any warnings appear, scrutinize to verify whether there are any dire warnings, otherwise the generated backup is safe to use. The only drawback in this scenario is the downtime impact, that is the time taken to back up the OCP clusters is longer than usual. Therefore, it is recommended that you back up the clusters during non-peak hours.

If the sync errors continue to appear, contact your IT department to check whether there is an issue with the storage infrastructure which might be preventing the sync command from completing on time.

Using DRS with CDW

You can back up and restore Kubernetes namespaces behind Cloudera Data Warehouse (CDW) entities (for example, Database Catalogs, Virtual Warehouses) on demand using the Data Recovery Service (DRS). CDW leverages DRS and provides CDP CLI endpoints which you can use to create and restore backups for CDW namespaces to back up CDW metadata and configurations such as Kubernetes objects, persistent volumes, autoscaling configuration, and so on.

The following limitations apply for CDW data service if you are on Embedded Container Service (ECS) or using an embedded database on Red Hat OpenShift Container Platform:

- The embedded database that CDW uses is part of the Control Plane. You cannot back up only CDW-related entities from the embedded database using the `dw create-backup` command. You must take a backup of the Control Plane service.
- You must restore the entire Control Plane configurations to restore configurations stored in the CDW database. This recreates the Control Plane namespace.

List of data recovery sub-commands for CDW

The following table lists the commands and CLI endpoints for backing up and restoring Kubernetes namespaces behind CDW entities:

DRS sub-commands for CDW	Description
<code>create-backup</code>	Creates an on-demand backup for the Data Warehouse including Kubernetes objects, persistent volumes, and so on. Backup requests are processed asynchronously and instantaneously.
<code>delete-backup</code>	Deletes an existing Data Warehouse backup. The call returns immediately. It returns a delete CRN, which is the deletion process identifier.
<code>describe-backup</code>	Returns the description of an existing Data Warehouse backup.
<code>restore-backup</code>	Restores the state of the Data Warehouse from an existing backup. It returns a restore CRN, which is the identifier of the restoration process.
<code>describe-restore</code>	Returns the description of the Data Warehouse restore operation.

DRS sub-commands for CDW	Description
list-backup-entities	Lists potential backup entities associated with the Data Warehouse.
list-backups	Lists backups associated with the Data Warehouse.
list-restores	Lists restores associated with the Data Warehouse.
get-logs	Returns the job logs corresponding to the specified CRN.

Related Information

[CDP CLI reference of DRS commands for CDW](#)