CDP Private Cloud Data Services Replication Manager

# Replication Manager Top Tasks

**Date published: 2023-05-24**
**Date modified: 2024-03-01**

## CLOUDERA

# Legal Notice

# Contents

# Use cases for Replication Manager in CDP Private Cloud Data Services

Replication Manager is a service in CDP Private Cloud Data Services. You can use this service to copy and migrate HDFS data, Hive external tables, and Ozone data between CDP Private Cloud Base 7.1.8 or higher clusters using Cloudera Manager version 7.7.3 or higher.

You can use Replication Manager for a variety of use cases. Some use cases include:

* Implementing a complete backup and restore solution.

  You might want to implement a backup and restore solution for HDFS data or Hive external tables. You create the replication policy based on the type of data you want to back up and restore. To implement this use case, you backup data in ClusterA to ClusterB. When the need arises, you can create another replication policy to restore the data from ClusterB to ClusterA.
* Replicating the required data to another cluster to run workloads or analytics.

  Sometimes, you might want to move workloads, especially heavy-duty workloads to another cluster to reduce the load and optimize the performance of the primary cluster, or run analytics on the required data on another cluster because you do not want to overload the primary cluster. In such scenarios, you can create replication policies in Replication Manager to replicate the required data at regular intervals. After replication, you can use the required tools to analyze the data.

## Replicating Ozone data in OBS buckets between secure clusters

Object store buckets (OBS) are storage buckets where all the keys are written into a flat namespace. It can be accessed using the S3 interface provided by Ozone. Ozone replication policies support data replication between OBS buckets in source and target clusters which support S3A filesystem using the S3A scheme or replication protocol. You can replicate Ozone data in OBS buckets between CDP Private Cloud Base 7.1.8 or higher clusters using Cloudera Manager version 7.7.3 or higher.

### Before you begin
The following task assumes that Auto-TLS and Kerberos are enabled on the source and target clusters, and Ranger is enabled on the source cluster.

### About this task
To replicate Ozone data between OBS buckets, perform the following tasks:

### Procedure

1. Login to CDP Private Cloud Data Services.
2. Add the source and target clusters to use in Ozone replication policy using the following steps:
   a) Ensure you have an *admin* or *poweruser* role.
   b) Go to the Replication Manager Clusters page, and click Add Cluster.

      The Management Console Clusters Cluster Details wizard appears.
   c) Enter the cluster details which include Cloudera Manager URL (use Cloudera Manager hostname if Auto-TLS is enabled), Username, Password, Location, and Data Center.
   d) Click Add.

      For more information, see Adding clusters to a CDP Private Cloud Data Services deployment
3. Complete the steps to configure SSL/TLS certificate exchange between the Cloudera Manager instances that manage the source and target clusters.

4. Complete the steps in additional configuration for OBS buckets to use the required OBS buckets in an Ozone replication policy.

5. Complete the steps in Preparing for Ozone replication policies to use the Kerberos-enabled clusters, and Ranger-enabled source cluster in the Ozone replication policy.

6. Complete the steps in Creating an Ozone replication policy to create and run the Ozone replication policy.

7. Click the replication policy on the **Replication Policies** page to view the job progress.

8. Verify whether the replicated data is available on the target cluster using the Ozone shell after the replication is complete.

# Tracking, monitoring, and troubleshooting replication policies

After you create a replication policy, you can track the replication policy job progress, monitor the status of the replication policies, and troubleshoot the failed replication policies in CDP Private Cloud Data Services Replication Manager.

### Tracking and monitoring replication policies

You can view, track, and monitor the available replication policies on the Replication Manager UI in CDP Private Cloud Data Services using the following pages:

- **Overview** page shows the statistics in the **Policies**, **Jobs**, and **Issues & Updates** panels.
- **Replication Policies** page lists all the replication policies and information about each replication policy.

### Troubleshooting replication policies

You can use one of the following methods to troubleshoot a failed replication policy:

- On the **Replication Policies** page, you can:
  - click Actions Collect diagnostic bundle for a replication policy to generate a diagnostic bundle for the replication policy. You can download the bundle as a ZIP file to your machine.
  - click Actions View Log for a replication policy to download, copy, or open the log to track the job and to troubleshoot any issues.

    **Tip:** On the Overview Issues & Updates panel, the Job Status column shows the current job status. If the job failed, click Failed to view the log details about the job.

- Click a job on the Replication Manager Replication Policies Job History panel to view the following job details:
  - Steps that Replication Manager ran for the job along with the timestamp and job details on the **General** tab.
  - Source and destination command output on the **Command Details** tab.
  - Stack trace for the commands that ran on the source Cloudera Manager for the failed job on the **Setup Error** tab.

  You can download the CSV reports from the General HDFS Replication Report field to track the replication jobs and to troubleshoot issues. The reports include performance statistics and file error listing.
- Click Cloudera Manager Running Commands for the source or target cluster. The recent command history shows the failed commands.
- Open the service logs in Cloudera Manager to track the errors on the source cluster and target cluster.

    **Tip:** You can also search on the Cloudera Manager Diagnostics Logs page to view the logs.

### Related Information
Managing HDFS replication policy
Managing Hive replication policy
Managing Ozone replication policy