Cloudera Replication Manager

Replication Manager Release Notes

Date published: 2019-11-15
Date modified: 2024-07-08

https://docs.cloudera.com/

Legal Notice

© Cloudera Inc. 2025. 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

What's new

This section lists major features and updates for Cloudera Replication Manager service.

Attention: Support of Replication Manager with Hortonworks Data Platform (HDP) clusters is Technical Preview and not available by default. For more information, contact your Cloudera account team. If you have questions regarding this service, contact support by logging a case on our Cloudera Support Portal.

May 2025

This release of the Replication Manager service introduces the following new features.

Retry the creation process for failed HBase replication policies

You can use the Actions Retry Failed Create option for HBase replication policies to retry the HBase replication policy creation process. This option appears for failed HBase replication policies that show the **Create failed** status. You can view the replication policy status on the Replication Manager Replication Policies page.

For more information, see Monitor HBase replication policy job details.

November 2024

This release of the Replication Manager service introduces the following new features.

CDP CLI commands for HBase replication policies

You can create, activate, or delete HBase replication policies using the following CDP CLI commands:

- activate-hbase-policy
- create-hbase-policy
- continue-hbase-setup
- delete-hbase-policy
- get-cluster-config
- list-paired-hbase-clusters
- suspend-hbase-policy
- update-hbase-policy
- verify-hbase-cluster-pair

For more information, see Creating HBase replication policy.

July 2024

This release of the Replication Manager service introduces the following new features.

Fine-grained permission to access CDP Public Cloud Replication Manager

You can choose to restrict user access to view and use CDP Public Cloud Replication Manager with RBAC entitlement. To enable or disable the role-based access control (RBAC) entitlement, contact your Cloudera account team.

For more information, see Fine-grained permission to access CDP Public Cloud Replication Manager.

Replication Manager support for AWS temporary credentials (technical preview)

You can use temporary AWS credentials, through the IDBroker service, to replicate HDFS data, Hive external tables, and HBase data from 7.1.9 SP1 Kerberized CDP Private Cloud Base clusters using Cloudera Manager 7.11.3 CHF7 or higher versions to CDP Public Cloud S3 clusters. You can also use the temporary AWS credentials to replicate HDFS data from S3 buckets to 7.1.9 SP1 Kerberized CDP Private Cloud Base clusters or higher using Cloudera Manager 7.11.3 CHF7 or higher versions.

This is a technical preview feature. It is not recommended for production deployments. Cloudera recommends that you try this feature in development or test environments. To enable this feature, contact your Cloudera account team.

For more information, see Add IDBroker to use temporary AWS session credentials

HBase replication policy enhancements

Specify custom username

Starting from CDP Public Cloud 7.2.18.200, you can specify a custom username in the Select Source Export snapshot user field during the HBase replication policy creation process. The option appears after you choose the Select Source Perform Initial Snapshot option. Replication Manager uses the specified username on the source cluster to export the initial snapshot to the target.

For more information, see Creating HBase replication policies.

Older releases

Overview of new features, enhancements, and changed behavior introduced in earlier releases of Replication Manager.

June 2024

This release of the Replication Manager service introduces the following new features.

HBase replication policy enhancements

Replicate data through an network load balancer (NLB)

During the HBase replication policy creation process, you can choose to specify an NLB if the source CDH 5.16.2 cluster uses the NLB to communicate with ZooKeeper and RegionServers of the destination Cloudera Manager of COD clusters.

Specify the NLB details after you enable the Select Destination Replicate via a Network Load Balancer option during the HBase replication policy creation process.

For more information, see Support matrix for CDP Replication Manager and Creating HBase replication policies.

Specify maximum number of tables to process in parallel

You can specify the maximum number of tables to process in parallel during the initial snapshot export and import step during an HBase replication policy run. You specify the number of tables using the Initial Snapshot Settings Maximum parallel snapshots option during the HBase replication policy creation process. If you do not enter any value, Replication Manager chooses an appropriate value, depending on the resources in the source and target cluster, to optimize the performance.

For more information, see Support matrix for CDP Replication Manager and Creating HBase replication policies.

April 2024

CDP version 7.2.18 introduces the following new features in the Replication Manager service.

Register GCP credentials

You can add the GCP credentials on the CDP Public Cloud Replication Manager Cloud Credentials page to use in Replication Manager.

For more information, see Registering GCS credentials to use in Replication Manager.

HBase replication policy enhancements

HBase replication policies support the following enhancements:

- You can create multiple HBase replication policies between multiple clusters to replicate HBase data.

  For more information, see Replicate HBase data simultaneously between multiple clusters.
- You can choose to Replicate all user tables or Replicate only tables where replication is already enabled after you choose the Select Source Replicate Database option during the HBase replication policy creation process.

  The Replicate only tables where replication is already enabled option is supported only if the target cluster uses Cloudera Manager version 7.12.0.0 and higher, 7.11.0-h3 and higher, or 7.9.0-h7 and higher.

  For more information, see Methods to replicate HBase data and Creating HBase replication policies.
- You can enter a YARN Queue Name to submit the replication job during the HBase replication policy creation or choose to retain the "default" value.

  For more information, see Creating HBase replication policies.
- You can click Actions on the **Replication Policies** page to Collect diagnostic bundle for the required HBase replication policy.

  For more information, see Monitor HBase replication policy job details.

For more information about feature support in CDP Public Cloud Replication Manager, see Supported features.

January 2024

This release of the Replication Manager service introduces the following new features.

Replicate HBase tables in a database depending on the replication scope

During the HBase replication policy creation process, after you select the Select Source Replicate Database , you can choose the Select Source Replicate all user tables option to replicate all the HBase tables in the database, or choose the Select Source Replicate only tables where replication is already enabled option to replicate the tables for which the replication scope is already set to 1. This provides a choice to replicate the required tables in a database.

The Replicate only tables where replication is already enabled option is supported only if the target cluster CDP version is CDP version 7.2.16.500 using Cloudera Manager 7.9.0-h7 or higher versions, CDP version 7.2.17.300 using Cloudera Manager 7.11.0-h3 or higher versions, or CDP version 7.2.18 using Cloudera Manager 7.12.0.0.

For more information, see Methods to replicate HBase data and Creating HBase replication policies.

Replicate HBase data between multiple clusters simultaneously

Starting from CDP Public Cloud version 7.2.16.500, 7.2.17.200, and 7.2.18 you can create multiple HBase replication policies between multiple clusters to replicate HBase data.

For more information, see Replicate HBase data simultaneously between multiple clusters.

August 2023

This release of the Replication Manager service introduces the following new features.

Replicate HBase data in all the tables in a database

You can replicate HBase data (existing tables and future tables in a database) using the Select Source Replicate Databases option during the HBase replication policy creation process. You can use this option only if the minimum target Cloudera Manager version is 7.11.0 and the minimum source cluster versions are CDH 5.16.2 (after you upgrade the source cluster Cloudera Manager), CDP Private Cloud Base 7.1.6, or COD 7.2.17.

For more information, see HBase replication policy and Creating HBase replication policies.

Provide a unique replication policy name

You must provide a unique name to the replication policy during the replication policy creation process if the Cloudera Manager API version is 51 or higher.

For other Cloudera Manager API versions, you can continue to use an existing replication policy with an empty name. However, if you edit the replication policy and provide a name for the replication policy, ensure that the name conforms to the validation rules.

For more information, see Managing HDFS replication policy, Managing Hive replication policy, or Managing HBase replication policy.

Load replication policies and their job history on the Replication Policies page

You can choose one of the following options to **Load policies faster by delaying to load their job history**:

- Delay loading job history when history is too long
- Always load job history
- Never load job history

By default, the replication policies are loaded only partially on the **Replication Policies** page, therefore the page might display incomplete statistics about a job status and replication policies with failed jobs might take a longer time to load. You can change the behavior depending on your requirements.

For more information, see Replication Policies page

June 2023

This release of the Replication Manager service introduces the following new features.

HBase replication policy enhancements

You can view the HBase RegionServer metrics for a specific replication peer in Replication Manager. You can use these metrics to monitor HBase replication jobs and to find and diagnose issues with the HBase replication peer.

For more information, see Graphical representation of HBase RegionServer replication peer metrics

CDP CLI commands for Replication Manager

You can collect and download the diagnostic bundle, add ABFS and AWS cloud credentials to use in Replication Manager and delete credentials using CDP CLI commands.

For more information, see Adding cloud credentials in Replication Manager using CDP CLI.

Mar 07, 2023

This release of the Replication Manager service introduces the following new features.

HBase replication policy enhancements

The following enhancements are available for HBase replication policies:

- During HBase replication policy creation, you can choose one of the following cloud credential options:
  - Credentials are available in source cluster HDFS service configuration setting- You can choose this option when you want to use a Google Cloud account. You can use this option for S3 and ADLS accounts as well.

    To add the advanced configuration settings for Google Cloud, Amazon S3, and ADLS accounts, perform the following steps:

    1. Go to the source Cloudera Manager Clusters *HDFS SERVICE* Configuration tab.
    2. Locate the HDFS Client Advanced Configuration Snippet (Safety Valve) for hdfs-site.xml property.
    3. Add the following key-value pairs to register a Google account to use in Replication Manager:
       - fs.gs.impl=com.google.cloud.hadoop.fs.gcs.GoogleHadoopFileSystem
       - fs.gs.project.id=*[***ENTER THE PROJECT ID***]*
       - fs.gs.system.bucket=*[***ENTER THE BUCKET NAME***]*
       - fs.gs.working.dir=/
       - fs.gs.auth.service.account.enable=true
       - fs.gs.auth.service.account.email=*[***ENTER THE SERVICE PRINCIPAL EMAIL ID***]*
       - fs.gs.auth.service.account.keyfile=/*[***ENTER THE LOCAL PATH OF THE P12 FILE***]*

       You can add these key-value pairs to register a Google account to use in Replication Manager from CDP Public Cloud 7.2.16.1 and higher versions.
    4. Add the following key-value pairs to register an S3 account to use in Replication Manager:
       - fs.s3a.access.key=*[***ENTER THE SESSION ACCESS KEY***]*
       - fs.s3a.secret.key=*[***ENTER THE SESSION SECRET KEY***]*
    5. Add the following key-value pairs to register an ADLS account to use in Replication Manager:
       - fs.azure.account.oauth2.client.id=*[***ENTER THE ABFS STORAGE CLIENT ID***]*
       - fs.azure.account.oauth2.client.secret=*[***ENTER THE ABFS STORAGE CLIENT SECRET KEY***]*
    6. Save and restart the HDFS service for the changes to take effect.
  - Credentials from External Account - You can choose this option for S3 and ADLS storage options.

  For more information, see Creating an HBase replication policy.
- You can Force Delete an HBase replication policy after Replication Manager fails to delete the replication policy (using the Actions Delete option on the **Replication Policies** page for the replication policy). For more information, see Managing HBase replication policy.

Jan 11, 2023

This release of the Replication Manager service introduces the following new features.

HBase replication policy enhancements

The following enhancements are available for HBase replication policies:

- HBase data replication using HBase replication policies is supported when SFT setting is automatically set on SFT-enabled clusters.

- You must choose the Select Destination I want to force the setup of this HBase replication policy option in the HBase policy creation wizard to acknowledge that the first-time setup between the selected source and destination clusters should be initiated after the existing pairing of the source and/or target cluster gets cleared.

  For more information, see Creating HBase replication policy.
- You can also use the following options to manage an HBase replication policy:
  - The Replication Policies HBase replication policy Actions Suspend action suspends all the active HBase replication policies between the source and target clusters selected in the replication policy.
  - The Replication Policies HBase replication policy Actions Activate action resumes data replication for all the HBase replication policies between the source and target clusters selected in the replication policy.
  - The Replication Policies HBase replication policy Actions Retry Create action retries the first-time setup between the source and target clusters selected in the replication policy. This option is available only if the first-time setup configuration fails.
  - The Replication Policies HBase replication policy Actions Retry Failed Snapshots action reruns the failed initial snapshots (and only the failed ones) in the replication policy if the Replication Manager failed to replicate the existing data of some tables.
  - The Replication Policies HBase replication policy Actions View Command Details action opens the latest replication policy job and shows the last 15 steps of the log for the policy job run. The steps and substeps appear in a tree view. The failed steps are expanded by default.

  For more information, see Managing HBase replication policy.
- You can now monitor the first-time setup configuration steps and its progress on the Replication Policies HBase replication policy Job History tab.

Hive replication policy enhancements

The following enhancements are available for Hive replication policies:

- You can use Hive replication policies to replicate Hive external tables and metadata from CDH 5.10 and higher clusters and CDP Private Cloud Base 7.1.1 and higher clusters to Data Hubs using Cloudera Manager 7.9.0 or higher.

  For more information, see Creating Hive replication policy.
- The Replication Policies Hive replication policy Actions View Command Details action opens the latest replication policy job and shows the last 15 steps of the log for the policy job run. The steps and substeps appear in a tree view. The failed steps are expanded by default.

  For more information, see Managing Hive replication policy.

May 16, 2022

This release of the Replication Manager service introduces the following new features.

Updates to HBase replication policy

- If you are using Cloudera Manager version 7.6.0 or higher, you can use the following options while creating an HBase replication policy:
  - Rolling HBase Service Restart on Source - This option appears if you select COD or Data Hub as the source cluster. Select this option to enable automatic rolling restart of HBase service on the source cluster after the HBase replication policy first-time setup steps are complete. Otherwise, Cloudera Manager performs a full restart of the service.
  - Rolling HBase Service Restart on Destination - Select this option to enable automatic rolling restart of HBase service on the target cluster as a rolling restart after the HBase replication policy first-time setup steps are complete. Otherwise, Cloudera Manager performs a full restart of the service.

- Validate Policy - Select to notify Replication Manager to verify the policy details after the policy creation is complete.

For more information, see Creating HBase replication policy.

- A warning message appears when you choose a cluster that is part of an existing cluster pairing. If there are HBase replication policies for the existing pairing, policy creation is not allowed to continue. If there are no policies, the existing pairing is removed and policy creation continues.

  For more information, see Using HBase replication policy.

Support for Replication Manager in ap-1 and eu-1 regional Control Planes

Replication Manager is now supported in the ap-1 (Australia) and eu-1 (Germany) regional Control Planes. For the list of all supported services in the supported Control Plane regions, see CDP Control Plane regions.

May 12, 2022

This release of the Replication Manager service introduces the following new features.

Migrating HDFS data from cloud storage to CDH clusters

After you register a CDH cluster as a classic cluster in Management Console, and register the cloud credentials (AWS S3 and Azure) in Replication Manager, you can migrate HDFS data from the cloud storage to the registered classic cluster (CDH cluster) using HDFS replication policies.

For more information about CDP CLI for HDFS and Hive replication policies, see Creating HDFS replication policy.

Tracking and monitoring the performance of HDFS and Hive replication policies

You can monitor and track the performance of the following replication policies:

- HDFS replication policies - Download the CSV reports in the HDFS Replication Report field on the Replication Policies *[***HDFS REPLICATION POLICY***]* Job History panel.
- Hive replication policies - Download the CSV reports in the HDFS Replication Report field and Hive Replication Report field on the Replication Policies *[***HIVE REPLICATION POLICY***]* Job History panel.

For more information, see Managing HDFS replication policy and Managing Hive replication policy.

Modifying HDFS and Hive replication policy options

After you create an HDFS replication policy or Hive replication policy, you can change the policy options as required to meet any changing requirements. Optionally, on the **Replication Policies** page, you can expand an HDFS or Hive replication policy to edit the policy description, frequency (start time cannot be modified if the policy has already started) to run the policy, YARN queue name to submit the replication job, maximum bandwidth for each map task, and maximum map slots or tasks per replication job. To optimize the replication policy performance, you can configure the queue name, maximum bandwidth, and maximum map slots as necessary.

For more information, see Managing HDFS replication policy and Managing Hive replication policy.

Restart HBase service when you use on-premises cluster as source cluster in HBase replication policy

After you create an HBase replication policy, you must restart the HBase service in the on-premises source cluster when the policy status on the **Replication Policies** page shows Manual restart (src) / restarting (dest) or Manual HBase restart needed on source. After the service restart is complete, the setup continues automatically for the replication policy.

However, if the source cluster Cloudera Manager version is 7.6.0 or lower and you are using an on-premises cluster as the source cluster, you must perform the following steps to complete the HBase replication policy setup:

1. Restart the HBase service on the on-premises source cluster when the policy status on the **Replication Policies** page shows **Waiting for 'Continue Setup' action call**.
2. Click Continue setup for the policy on the **Replication Policies** page after the service restart is complete. This action informs Replication Manager to continue the replication policy setup.

For more information, see Managing HBase replication policy.

Mar 1, 2022

This release of the Replication Manager service introduces the following new features.

Using CDP CLI for HDFS and Hive replication policies

The CDP CLI commands for Replication Manager are under the replicationmanager CDP CLI option.

You can create, suspend, activate, or delete HDFS and Hive replication policies using create-policy, suspend-policy, activate-policy, and delete-policy CDP CLI commands. You can also list the clusters, replication policies, service statuses, credentials, and get the credentials for a specific cluster using the list-clusters, list-policies, list-cluster-service-statuses, list-all-credentials, and get-credentials CDP CLI commands.

For more information about CDP CLI for HDFS and Hive replication policies, see CDP CLI for Replication Manager.

Feb 24, 2022

This release of the Replication Manager service introduces the following new features.

Creating HDFS and HBase snapshot policies (Tech preview)

A snapshot is a set of metadata information, a point-in-time backup of HDFS files and HBase tables. You can create snapshot policies for HDFS directories and HBase tables in registered classic clusters and SDX Data Lake clusters to take snapshots at regular intervals. Before you create an HDFS snapshot policy for an HDFS directory, you must enable snapshots for the directory in Cloudera Manager.

After a snapshot policy takes a snapshot of the HDFS directory or HBase table, you can perform the following tasks:

- Restore the snapshot to the original directory or table using the Restore Snapshot option.
- Restore a directory or table to a different directory or table using the Restore Snapshot as option.

For more information about snapshot policies, see Snapshot policies in Replication Manager.

Important: Creating an HDFS or HBase snapshot policy is a technical preview feature. Access to preview features is provided upon request to customers for trial and evaluation. The components are provided 'as is' without warranty or support. Further, Cloudera assumes no liability for the use of preview components, which should be used by customers at their own risk. Contact your Cloudera account team to have this preview feature enabled in your CDP account.

Override default storage location for replicated Hive external tables in the target cluster

Administrators can override the default storage location for replicated Hive external tables in the target cluster when they create a Hive replication policy.

Before you add another path to override the default storage location, ensure that the following steps are complete in the Ranger UI:

1. Alter the ranger policy Default: Hive warehouse locations in cm_s3 service to allow the Hive service to access the updated locations of S3 bucket path.
2. Manually update the Ranger and Sentry permissions.

For more information about editing the Hive External Table Base Directory option in Hive replication policies, see Creating Hive replication policy.

Cannot suspend HBase replication policies

If you create an HBase replication policy, you can no longer suspend the policy. However, you can resume any existing suspended HBase replication policy.

Generate and download diagnostic bundles for replication policies

You can generate and download diagnostic bundles for HDFS and Hive replication policies. You can use the bundle to troubleshoot failed replication jobs or to view replication-specific diagnostic data for a replication policy.

For more information about CDP CLI for HDFS and Hive replication policies, see Managing HDFS replication policies and Managing Hive replication policies.

Sept 9, 2021

This release of the Replication Manager service introduces the following new features.

Enhancements to HBase replication policies

The following enhancements are available for an HBase replication policy:

- You do not need to create a schema similar to the source cluster on the destination cluster.
- Replication Manager performs the first-time setup configuration steps which includes HBase service restart on both the clusters automatically.
- Optionally, you can also create or use an existing HBase replication machine user during policy creation and then validate the existing username with the UMS. The username and password is automatically synchronized to the destination cluster's environment (and to the source's as well if the source is Cloudera Operational Database (COD)).

April 30, 2021

This release of the Replication Manager service introduces the following new features.

Registering CDP Private Cloud Base clusters in CDP Public Cloud

You can register the CDP Private Cloud Base as a classic cluster in CDP. After registration, you can replicate the data in HDFS and Hive external tables in the classic cluster to CDP Public Cloud.

March 25, 2021

This release of the Replication Manager service introduces the following new features.

Replication Manager introduces HBase replication policy which supports the following use cases:

- From CDP Private Cloud Base cluster to Data Hub cluster.
- From CDH cluster to Data Hub cluster.
- From CDH cluster to COD cluster.
- From COD cluster to COD cluster.

Important: HBase replication policy is a technical preview feature. If you require access, contact your Cloudera account team.

June 19, 2020

This release of the Replication Manager service introduces the following new features.

Replication Manager is a service for copying and migrating data between environments within the enterprise data cloud. It is a simple, easy-to-use, and feature-rich data movement capability to move existing data and metadata to the cloud to fuel new workloads.

Replication Manager introduces the following new features:

- Availability of Azure Blob Filesystem (ABFS) cloud storage
- Ability to use Cloudera Data Platform Data Center clusters

November 15, 2019

This release of the Replication Manager service introduces the following new features.

Replication Manager is a service for copying and migrating data between environments within the enterprise data cloud. It is a simple, easy-to-use, and feature-rich data movement capability to move existing data and metadata to the cloud to fuel new workloads.

Replication Manager introduces the following new features:

- Full support for CDH On-Premise cluster (Classic Cluster)
- AutoTLS enablement
- HDFS-less compute clusters
- Table Level Replication

Known Issues

This section lists known issues that you might run into while using the Replication Manager service.

DOCS-13504

When you create an HBase replication policy between a source cluster using Cloudera Manager version is 7.6.0-patch5366 or higher and a target cluster using Cloudera Manager version is 7.6.0 or lower, the first-time setup is not initiated and the following misleading message appears during policy creation:

Skipping Replication Setup because it has already been done.

To resolve this issue, upgrade the target cluster to Cloudera Manager version to 7.6.0-patch5366 or higher.

DMX-518

Hive replication policy fails when a table is dropped in the source database (export tables) when the replication job is running.

DMX-519

If snapshots on the /user/hive/warehouse directory is not enabled, Hive replication policy fails when inserts are done on a source table during replication job run.

DMX-521

While running Hive replication policy, if you drop a table, it is not dropped on the target cluster. The data still remains and 'show tables' displays the dropped table after successful replication instances.

DMX-1455

An HBase replication policy from Cloudera Operational Database COD to COD cluster fails if you select the Perform Initial Snapshot option in the Create Replication Policy wizard, and the source and destination COD clusters use different AWS accounts.

DMX-2923

When you create an HBase replication policy with the Perform Initial Snapshot option in a bi-directional replication setup from the disaster recovery cluster to the production cluster, the original table gets dropped and it is recreated with the content in the disaster recovery cluster. Therefore, the original content of the table is lost.

When you create the HBase replication policy in a bi-directional replication setup from the disaster recovery cluster to the production cluster, do not select the Perform Initial Snapshot option.

OPSAPS-66327

The HBase peer can be created in the disabled state in CDH 6.0 or higher clusters. However, HBase peer cannot be created in disabled state in CDH 5 clusters and the peer remains enabled throughout the replication setup process. As a workaround, you can create a 'test' HBase replication policy with a small table in CDH 5 clusters. This creates the HBase peer.

OPSAPS-69210

HBase replication policies do not support source or destination CDP Public Cloud clusters created using the Micro Duty template.

OPSAPS-69782

HBase replication using HBase replication policies between two Data Hubs/COD clusters fail if all the following conditions are true:

- The destination Data Hub/COD cluster's Cloudera Manager version is 7.9.0-h7 through 7.9.0-h9 or 7.11.0-h2 through 7.11.0-h4, or 7.12.0.0.
- The source Data Hub/COD cluster's Cloudera Manager major version is higher than the destination cluster's Cloudera Manager major version.
- The Initial Snapshot option is chosen during the HBase replication policy creation process and/or the source cluster is already participating in another HBase replication setup as a source or destination with a third cluster.

Upgrade the destination cluster's Cloudera Manager to a major version that is equal to or greater than the source cluster's Cloudera Manager major version.

OPSAPS-62230

HBase replication command steps for an HBase replication policy might fail if you have configured an automatic delete of /tmp folder content on the destination cluster's Cloudera Manager server node.

OPSAPS-70282

You cannot create HBase replication policies if the target Cloudera Manager Clusters HBase service Configuration HBase ZooKeeper Secure Client Enabled property is enabled.

Disable the HBase ZooKeeper Secure Client Enabled property, and then create HBase replication policies.

OPSAPS-71317, OPSAPS-71318

When you delete an HBase replication policy, its peer is not removed if a table name in the replication policy contains a period. This issue is observed when the source Cloudera Manager is 7.4.4, 7.6.1, 7.6.7, 7.7.1, 7.7.3 CHFx, or CDP Public Cloud 7.2.16 using source Cloudera Manager 7.9.2, or CDP Public Cloud 7.2.17 using source Cloudera Manager 7.11.0, and unsupported target Cloudera Manager versions.

To resolve the issue:

- Upgrade to a supported target Cloudera Manager.
- Upgrade the source Cloudera Manager to 7.11.3 CHFx or higher.
- Remove the period from the table name on the source cluster.

- Perform the following steps to remove the peer associated with the deleted replication policy manually:
  1. Confirm that the HBase peer in the source HBase cluster does not contain any tables. If there are any tables, this indicates that there might be other HBase replication policies between these clusters. In such instances, you can contact your Cloudera account team.
  2. Manually remove the peer using the remove_peer *[***PEER ID***]* HBase command.
  3. Create an HBase replication policy in Replication Manager.

DMX-3888

The existing Hive replication policies do not appear on the **Replication Policies** page after you upgrade the target cluster from CDP Public Cloud 7.2.16 using Cloudera Manager 7.9.0 to CDP Public Cloud 7.2.17 using Cloudera Manager 7.11.0-h2. This issue appears if the existing replication policies were created using external APIs and the API payload included the rangerImport Properties and sentryExportProperties attributes.

Before you upgrade the target cluster, verify whether the HIVE_REPLICATION_SENTRY_EXPORT_PROPERTIES feature flag is enabled. If the feature flag is not enabled, you must remove the rangerImportProperties and sentryExportProperties attributes from the existing Hive replication policies, and then upgrade the target cluster.

Related Information

Replication Manager Overview

Fixed Issues

This section lists fixed issues in the Replication Manager service.

CDPSDX-2879: Ranger import fails when you create a Hive replication policy for a medium duty Data Lake cluster

When you create a Hive replication policy with the Include Sentry Permissions with Metadata or Skip URI Privileges option for a medium duty Data Lake cluster, Ranger import fails. Before you choose the Include Sentry Permissions with Metadata option for a Hive replication policy for a medium duty Data Lake cluster, contact Cloudera Support.

DMX-348: Export remote Hive Metastore step failed while using Replication Manager

While creating a replication policy with no Sentry permissions for the source database/tables, with latest Data Lake clusters, an error message appears The remote command failed with error message : Command (HDFS replication (433)) has failed.

Make sure that while using CDH cluster(s), you do not use MapReduce 1 service.

DMX-355: Currently, Replication Manager does not work efficiently with high-frequency policies

Recommend using policy frequency which is greater than 30 minutes.

DMX-364: Multiple instance of Replication Manager UI displays as "In Progress"

When the replication policy is scheduled with a specific interval, it is seen from the UI that all the instances are "In Progress" and one of the instances is failed.

Refresh the web browser and try again.

DMX-391: Schedule Run option displays only "Does not repeat" and "Custom" options

Recommend to use "Does not repeat" or "Run now" option to schedule the replication policy.

DMX-489

With default replication configurations, small files consume more time to complete the replication process. You can increase the number of mappers to improve the replication performance.

DMX-636: Inconsistency in the value of Timestamp Type column when source ORC is replicated, and source and target clusters are in different time zones

From an Auto TLS source cluster, when a ORC table with static data is replicated, the data in the Timestamp Type column does not match in the target cluster.

DMX-666: Replication fails when the exception "connection timed out" is not handled in Cloudera Manager

Ensure there is connectivity between Source Cloudera Manager and SDX CM. If the source hostname is not resolved to IP, add the host mapping of Cloudera Manager host to /etc/hosts entries of SDX CM.

OPSAPS-61288: Replication setup fails when a non-default namespace is not created on the destination cluster

Create a namespace on the destination cluster before you create an HBase replication policy.

OPSAPS-61596: HBase policy returns "different schema" when tables on source and destination clusters have the same column families

This issue appears because HBase replication policies handle tables that have table attributes incorrectly. Tables with table attributes in replicated tables might lead to other errors as well.

Remove the table attributes for the HBase tables.

OPSAPS-62836

When the first HBase replication policy is created between two clusters where the source cluster is an on-premises cluster, sometimes the policy's status shows Waiting for 'Continue Setup' action call for the first few seconds. No user action is required. The status automatically is updated to Configuring clusters after a few seconds.

OPSAPS-62910, OPSAPS-62948, OPSAPS-62956, and OPSAPS-62998

When you perform HBase replication policy creation, policy update, or policy delete operations on multiple policies between the same cluster pair at once, different failures appear. This is because the HBase peer does not get synchronized during these operations as expected.

Perform HBase replication policy creation, modification, or deletion action one at a time on each HBase replication policy of a cluster pair.

OPSAPS-62995

The HBase replication policy first time setup fails when the destination cluster is a Data Hub.

This issue appears when the HBase classpath is not configured automatically as expected.

To resolve this issue, perform the following steps in Cloudera Manager of the Data Hub before setting up an HBase replication policy:

1. Navigate to the HBase service.
2. Search for the following `Advanced Configuration Snippet (Safety Valve)` values, and add the key and value as shown below:
   - `HBase Service Environment Advanced Configuration Snippet (Safety Valve)`

     Key: HBASE_CLASSPATH

     Value: HBASE_CLASSPATH="$HBASE_CLASSPATH:/opt/cloudera/parcels/CDH/lib/hbase-cdp-jars/"
   - `HBase Client Environment Advanced Configuration Snippet (Safety Valve)`

     Key: HBASE_CLASSPATH

     Value: HBASE_CLASSPATH="$HBASE_CLASSPATH:/opt/cloudera/parcels/CDH/lib/hbase-cdp-jars/"
3. Restart HBase service.

OPSAPS-63071

HBase replication policies from on-premises (CDH5 and CDH6) clusters fail when the source cluster Cloudera Manager version is 7.6.0.

To resolve this issue, perform the following steps in the source cluster Cloudera Manager before you create an HBase replication policy:

1. Navigate to the HBase service.
2. Search for the following `Advanced Configuration Snippet (Safety Valve)` values, and add the key-value pairs as shown below:
   - `HBase Service Environment Advanced Configuration Snippet (Safety Valve)`

     Key: HBASE_CLASSPATH

     Value: HBASE_CLASSPATH="$HBASE_CLASSPATH:*[***LOCATION OF HBASE REPLICATION PARCEL JAR FILE***]*"
   - `HBase Client Environment Advanced Configuration Snippet (Safety Valve)`

     Key: HBASE_CLASSPATH

     Value: HBASE_CLASSPATH="$HBASE_CLASSPATH:*[***LOCATION OF HBASE REPLICATION PARCEL JAR FILE***]*"

   For example, the location of the HBase replication parcel jar in your CDH source cluster might be /opt/cloudera/parcels/CLOUDERA_OPDB_REPLICATION-1.0-1.CLOUDERA_OPDB_REPLICATION5.14.4.p0.9261092/lib/
3. Restart HBase service.

OPSAPS-63138

The `HBase Replication First Time Setup` command runs successfully in the destination cluster though the `Admin Setup HBase replication` subcommand fails on the source cluster.

OPSAPS-63905

Replication Manager runs the clone_snasphot command when restoring snapshots on SFT-enabled clusters without setting the SFT attributes on the table.

OPSAPS-64034

When you delete an HBase replication policy, the HBase peer used in the policy is also deleted if a replicated table name has the - character. Note that the HBase peer is deleted even when there are existing HBase replication policies using this peer. This issue also occurs if a replicated table in any of the policies that uses this peer has the - character in its name.

OPSAPS-64879

Replication policies with an empty name do not appear on the Replication Policies page.Provide a unique replication policy name during replication policy creation.

OPSAPS-65572

When you create an HBase replication policy between a cluster pair where HBase policies between them are suspended (which also means that the corresponding HBase peer is disabled), the HBase peer is enabled at the end of the policy creation process which results in an inconsistent suspend state for the suspended policies.

OPSAPS-65573

HBase replication policy creation and the Retry Failed Snapshots option can be run concurrently.

OPSAPS-66303

A NullPointerException appears when you delete an HBase replication policy for which the policy creation failed because a table in the policy did not exist on the source cluster. After the failed delete

action (using the Actions Delete option on the **Replication Policies** page), select Actions Force Delete to get the policy deleted.

OPSAPS-66305

If the snapshot export/import fails during HBase replication policy creation, and then you choose the Actions Retry Failed Snapshots option for the replication policy on the **Replication Policies** page, the HBase peer is not enabled even if there were no export/import failures during the retry action. In this scenario, the replication peer in the HBase service remains disabled and the Retry Failed Snapshots action fails.

Delete the replication policy and re-create it.

OPSAPS-66327

The HBase peer can be created in the disabled state in CDH 6.0 or higher clusters. However, HBase peer cannot be created in disabled state in CDH 5 clusters and the peer remains enabled throughout the replication setup process. As a workaround, you can create a 'test' HBase replication policy with a small table in CDH 5 clusters. This creates the HBase peer. This issue is fixed.

OPSAPS-66924

When the snapshot export fails during the HBase replication policy job run, the target HBase folder in the destination Data Hub or COD gets deleted if you are using Cloudera Manager versions that are lower than 7.6.7 CHF8, 7.11.0, or 7.9.0-h6 on the source cluster.

As a workaround, you can either revoke the delete permission for the user, or ensure that you use an access key/role that does not have delete permissions for the required storage component. For more information on how to create an access key in AWS or Azure service principal, see Target HBase folder is deleted when HBase replication policy fails.

OPSAPS-66988

You cannot create HBase replication policies if the target CDP version is 7.2.16, 7.2.16.1, 7.2.16.2, or 7.2.16.3 and the source Cloudera Manager version is 7.7.3 or lower, and the source Cloudera Manager API version is v50 or higher.

Tip: The endpoint http://*[***CM_HOST***]*:*[***CM_PORT***]*/api/version shows the API version of the Cloudera Manager.

OPSAPS-67042

Currently, when you create an HBase replication policy for a table, a snapshot is created and then the replication scope is set to "1". Because of the order of steps, the data that is generated between snapshot creation and setting the replication scope is not replicated.

As a workaround, you can set the replication scope of the tables on the source cluster that you want to replicate to "1" and then create the HBase replication policy.

To configure the replication scope for a table on the master cluster, run the "alter *[***TABLE_NAME***]*, {NAME => *[***COLUMN-FAMILY***]*, REPLICATION_SCOPE => 1} command for each column family that must be replicated. `REPLICATION_SCOPE` is a column-family level attribute, where the value '0' means replication is disabled, and '1' means replication is enabled.

OPSAPS-70074

When you choose the Advanced Force Overwrite option during the Hive replication policy creation process to replicate data from CDP Private Cloud Base cluster to CDP Public Cloud, the `Check Hive Metadata` step is skipped during the replication policy run.

OPSAPS-62230

HBase replication no longer fails if the content of the /tmp folder gets deleted while Cloudera Manager is running.

OPSAPS-70297

You can specify a custom username in the Select Source Export snapshot user field during the HBase replication policy creation process. The option appears after you choose the Select Source Perform Initial Snapshot option. Replication Manager uses the specified username on the source cluster to export the initial snapshot to the target.