NetBackup™ Web UI Apache Cassandra Administrator's Guide

Release 10.3

VERITAS™

NetBackup™ Web UI Apache Cassandra Administrator's Guide

Last updated: 2023-10-26

Legal Notice

Veritas Technologies LLC
2625 Augustine Drive
Santa Clara, CA 95054

http://www.veritas.com

Technical Support

Technical Support maintains support centers globally. All support services will be delivered in accordance with your support agreement and the then-current enterprise technical support policies. For information about our support offerings and how to contact Technical Support, visit our website:

https://www.veritas.com/support

You can manage your Veritas account information at the following URL:

https://my.veritas.com

If you have questions regarding an existing support agreement, please email the support agreement administration team for your region as follows:

Worldwide (except Japan)	CustomerCare@veritas.com
Japan	CustomerCare_Japan@veritas.com

Documentation

Make sure that you have the current version of the documentation. Each document displays the date of the last update on page 2. The latest documentation is available on the Veritas website:

https://sort.veritas.com/documents

Documentation feedback

Your feedback is important to us. Suggest improvements or report errors or omissions to the documentation. Include the document title, document version, chapter title, and section title of the text on which you are reporting. Send feedback to:

NB.docs@veritas.com

You can also see documentation information or ask a question on the Veritas community site:

http://www.veritas.com/community/

Veritas Services and Operations Readiness Tools (SORT)

Veritas Services and Operations Readiness Tools (SORT) is a website that provides information and tools to automate and simplify certain time-consuming administrative tasks. Depending on the product, SORT helps you prepare for installations and upgrades, identify risks in your datacenters, and improve operational efficiency. To see what services and tools SORT provides for your product, see the data sheet:

https://sort.veritas.com/data/support/SORT_Data_Sheet.pdf

Contents

Chapter 1

Overview

This chapter includes the following topics:

Overview of NetBackup for Apache Cassandra

Table 1-1 Steps to configure and protect Apache Cassandra Cluster

Step	Action	Description
Step 1	Sign in to NetBackup web UI as the Default Security Administrator. Then add the Apache Cassandra user to the **Default Apache Cassandra Administrator** role.	See the *Default Apache Cassandra Administrator* role in *NetBackup web UI Administrator's Guide*. **Note:** The **Default Apache Cassandra Administrator** role has the minimum required permissions to perform the Apache Cassandra administrator tasks.
Step 2	Add and manage credentials.	See "Add DSS Clusters" on page 18.
Step 3	Protect Cassandra Clusters.	See "Protecting Apache Cassandra assets " on page 26.

Table 1-1 Steps to configure and protect Apache Cassandra Cluster *(continued)*

Step	Action	Description
Step 4	Recover Cassandra cluster, key spaces or column family.	See "Recovering Apache Cassandra assets" on page 30.

NetBackup Apache Cassandra support overview

Apache Cassandra is a popular scale-out NoSQL database. Cassandra runs on commodity hardware with direct-attached storage. A typical Cassandra cluster consists of nodes that store data. Cassandra replicates data among the nodes to provide resiliency against node downtimes. There is no notion of a primary copy of data and any node may have a more recent version of data record than its replicas. One of the important characteristics of Cassandra is that it prefers availability over consistency. The database is always available even if the replicas of data are not always up to date.

NetBackup Cassandra Protection

NetBackup provides advanced solution to protecting Cassandra clusters. The solution has the following characteristics:

1. **Agentless**: No need to place backup agents on Cassandra cluster nodes. Effectively, there is no code that hinders high-performance Cassandra cluster.
2. **Single pass data copy:** During backup, a thin client is used to make a single pass over the Cassandra data files (called sstables) to minimize IO footprint.
3. **Off-host data optimization:** Cassandra data is replicated for resiliency. Backups are for longer retention. NetBackup Cassandra solution processes data to:
   - Determine a cluster-consistent point-in-time.
   - Remove replica records.
   - Remove stale data that caused by record overwrites.

   All this processing happens off-host on Data staging servers to ensure that backup processes do not affect your high-performance Cassandra clusters.
4. **Incremental backups:** NetBackup supports incremental backups of Cassandra to optimize backup times after a full backup. The solution automatically detects new key spaces or column families to take a full backup of these new structures while incremental backups of previously existing structures perform.

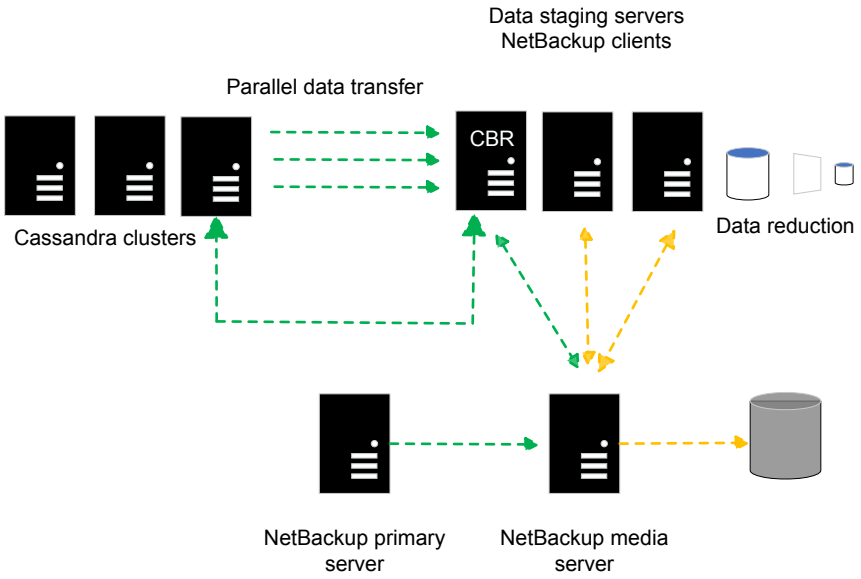
5. **Scalable Backup:** Cassandra lets you easily scale your Cassandra cluster by adding more nodes whenever required. It automatically redistributes the existing data to new nodes while the cluster is online. NetBackup Cassandra protection is scalable and lets you add more Data Staging Servers to meet your backup requirements.
6. **DataCenter Identification:** NetBackup Cassandra protection can be configured to backup data from a specific datacenter. It queries Cassandra cluster and automatically identifies the nodes present in various datacenters. It then engages only the nodes in the specific datacenter for backing up the data.
7. **DataCenter aware restore:** At the time of restore, NetBackup connects to the restore cluster and determines its current topology. The solution reconciles this topology with the one present at the backup time to allow for changes in the topology and restores considering the current topology. The solution provides more options for changing the datacenters, number of replicas in each datacenter, change in keyspace and column family names, etc. to help you with your restore requirements.

   Note: Indexes are not restored in case of rename column family scenario. So you add an index to the renamed column family.

8. **Granular restore:** NetBackup Cassandra solution lets you restore a part of the backup data set. You have option to restore a few of the key spaces or only some of the column families.
9. **Repair-less Restore:** The restore processes ensure that after data is restored, there is no need to perform further recovery steps. The data is available immediately after a restore in your high-performance Cassandra cluster.

NetBackup Apache Cassandra Protection Architecture

Figure 1-1 Architecture overview



In this architecture:

- NetBackup primary server have backup policies and schedules. It is responsible for managing backup jobs.
- NetBackup media server have the backup data. All NetBackup backup targets are supported for Cassandra protection.
- Data staging servers perform off-host processing of Cassandra data to:
  - Determine a cluster-consistent point-in-time.
  - Remove replica records.
  - Remove stale data caused by record overwrites.
- To perform off-host processing, the data staging server must have Cassandra installed on these nodes. NetBackup expects a Cassandra cluster of the same

distribution and version configured on the data staging server. If you have an SSL based authentication and/or LDAP configuration on the DataStax application cluster, then a same configuration of authentication must be performed on the data staging servers with the same root CA certificate as the application cluster. Maintain the version of Cassandra on data staging servers as you do for the Cassandra clusters.

- One of the nodes in data staging servers is setup as a “CBR” node (Cassandra Backup and Restore node). CBR performs the entire orchestration required for performing an effective backup and restore.

During backup, the production data is copied to the data staging servers. The data is then deduplicated and transfered to the backup hosts / NetBackup media servers. One data stream is written per DSS. If you have multiple DSS nodes, then data is streamed parallely or concurrently from these DSS nodes. NetBackup recommends to have the same number of streams configured on the backup hosts collectively to get maximum performance. Hence, number of streams per backup host × number of backup hosts >= Data staging servers.

During restore, the data is staged onto the data staging servers from the NetBackup media servers. This staged data is then restored into the Cassandra production cluster as per the number of replicas and Data centers configured for the keyspace being restored.

During restore, you can choose to:

- Restore the entire Cassandra cluster.
- Restore some keyspaces and/or column families.
- Rename some keyspaces and/or column families.
- Reconfigure data replica for the data that would be restored.

Pre-requisites and best practices

- Ensure that NetBackup supports the installed Cassandra version. For more information, refer Software Compatibility List.
- Backup host, Data staging server and Cassandra are supported on RHEL platform only.
- NetBackup requires the same distribution Apache/DataStax and version on the Data Staging Server (DSS) cluster as the production cluster which is being protected.
  NetBackup supports *yum* and *tar-based* deployments for DataStax Cassandra cluster in DSS and in production. DSS and production clusters must have same type of deployments.

- NetBackup requires around 20% of the nodes of the datacenter being protected as DSS.
- The DSS should be added to the backup environment so that NetBackup can perform the following:
  - Stage the data to the DSS.
  - Deduplicate the data saving to the backup storage.
  - Copy the data to NetBackup media.
- The DSS should have the same version of Cassandra as the Cassandra production cluster.
- The DSS and Cassandra production credentials must be added in NetBackup credential management console before adding DSS and Cassandra production clusters. Then on the DSS and Cassandra production cluster add workflow, must select desired credentials from the existing credential list.
- NetBackup supports SSL, LDAP and DataStax Cassandra with simple authentication. Use database username and password to connect Cassandra and to run commands like `cqlsh` and `nodetool utils`. Configure Cassandra in the NetBackup credentials during DSS cluster configuration and Cassandra cluster configuration.
- Enable SSH on all the Cassandra nodes and DSS nodes.
- Ensure that the local time of Cassandra nodes, the DSS, and the backup hosts are synchronized with NTP server.
- Configure a non-root host user account for the data staging server cluster in NetBackup credentials management.

  Note: The non-root host user account can be separate or the same. It must be valid with a home folder and rights to connect to the respective nodes with a use of `ssh`. Add the host user in the `sudoers` file on the respective nodes. Database username and password must be same on DSS and application cluster.

- Before you run Cassandra backup or restore, ensure that you received a successful ping response from all the data staging servers to Cassandra nodes and the backup host.
- Check and update the firewall settings for the backup hosts, data staging servers, and Cassandra nodes can communicate.
- Ensure that the specified paths in the DSS cluster configuration are existing on all the DSS and Cassandra nodes.

- Whenever you upgrade Cassandra or make any schema change like delete a keyspace or column family, initiate a full backup before any incremental backup job.
- Ensure that the specified host user account for the cluster has read and write access to the specified folders in the DSS cluster configuration.
- Host mapping must be done according to the IP preference.
- Ensure that SStableloader utility works between the production nodes and data staging server.
- Ensure that free space and the memory on the DSS is three times larger than the column family in the Cassandra cluster. Maintain similar memory size on all the DSS nodes.

  Note: The compaction operation on the DSS needs more memory. Deploying higher RAM on the DSS nodes will result in better backup and restore performance.

- Maintain a minimum 20% free space on Cassandra nodes during backup operations.
- Ensure enough free space on target cluster nodes during the restore as per the size of data being restored.
- Before the restore, ensure that the target Cassandra version has the same version as the version you backed up from.
- Before the restore ensure that the target cluster and target Data Staging Server cluster are fully configured in NetBackup.
- Canceling a parent job in a compound restore job does not cancel the child restore jobs. You must manually cancel the child restore jobs.
- Ensure that Connections per host (cph) value is set to 1 in DSS settings for Datastax cassandra backup.

RBAC permissions for a Cassandra role

- Ensure to assign both create and update permissions to:
  - Add DSS cluster.
  - Add Apache Cassandra cluster.
  - Add DSS nodes.
  - Edit Apache Cassandra cluster.

- The database credentials of DSS cluster should be the same as Cassandra production cluster.
- You must disable the `requiretty` option globally in the `sudoers` file, by replacing `Defaults requiretty` with **`Defaults !requiretty`**.

Note: This action changes the global `sudo` configuration.

- In case of `tarball` based installation, you must always start Cassandra services from `tarball installation` bin path location.
- For database user account, if `default_scheme` is **internal** for `authentication_options` in `dse.yaml` file, then specify the internal authentication user. If `default_scheme` is set to **LDAP**, then specify the LDAP user account.
- For NetBackup versions upgraded from versions prior to 10.2.1, you need to trigger the discovery manually for both DSS and production cluster.
- The database user account configured in NetBackup for the following must have all the required permissions in the cluster:
  - DSS cluster
  - Backup and restore of Cassandra production cluster.

  The user must be able to Create, View, Update and Drop any resources in the cluster. On the DSS cluster you can provide specific permissions or assign the superuser role to the configured database user account.
- Ensure that the DSS distribution, working directory and script home directory paths under Cassandra configuration are not the same.

Note: Working directory path cannot be set as `/root`.

- Ensure that you update the `secure_path` list with Java executable path in `/etc/sudoers` file.
- Modify the `cassandra.yaml` file to set the following parameters on all DSS nodes:

Parameters	Description/Value
cluster_name	Name of the cluster. **cluster_name: <Provide name of DSS cluster>**

Parameters	Description/Value
num_tokens	Set num_tokens as 1. **num_tokens: 1**

Initial_token	Calculate and set Initial_token using the following command:
	`python -c "print [str(((2**64 / number_of_nodes_in_cluster) * i) - 2**63) for i in range(number_of_nodes_in_cluster)]"`
	`initial_token: <To be calculated>`

Incremental backups	Disable incremental backups. **Incremental backups: false**
snapshot_before_compaction	Disables taking a snapshot before each compaction. **snapshot_before_compaction: false**
auto_snapshot	Disable auto snapshot. **auto_snapshot: false**
compaction_throughput_mb_per_sec	Disable compaction throttling. **compaction_throughput_mb_per_sec: 0**
hinted_handoff_enabled	Disable hinted handoff. **hinted_handoff_enabled: false**
cdc_enabled	Disable CDC functionality. **cdc_enabled: false**
enable_user_defined_functions	Enable user-defined functions. **enable_user_defined_functions: true**
enable_scripted_user_defined_functions	Enable scripted user-defined functions. **enable_scripted_user_defined_functions: true**

Components and Terminologies of the Cassandra backup and recovery

The following table describes the purpose of different components and terminologies of the Cassandra backup and recovery solution.

Table 1-2

Components and Terminologies	Purpose and Definition
Application Cluster	■ Application cluster is the Cassandra production cluster name. ■ Cluster name must be a single word with no white spaces in between words and must be the actual cluster name used in the `Cassandra.yaml` file on the production nodes.
Protection plan	A protection plan defines when backups are performed, how long the backups are retained, and the type of storage to use. Once a protection plan is set up, assets can be subscribed to the protection plan.
Backup host	The backup host acts as a proxy client. All the backup and the restore operations are executed through the backup host. The Cassandra Backup Recovery (CBR) solution, uses the BigData policy with application type **cassandra**. The Protection plan uses this backup host. The media server that is used to configure storage server for the CBR solution must be used as backup host. **Note:** You can also use NetBackup client as a backup host.
Cassandra cluster	Represents the Cassandra production cluster that you want to protect
Cassandra Backup Recovery component	The NetBackup thin client which gets deployed on data staging servers and Cassandra cluster to aid in backup and restore operations.

Table 1-2 *(continued)*

Components and Terminologies	Purpose and Definition
Data staging servers	NetBackup requires a set of servers for backup of Cassandra cluster in addition to the NetBackup primary, and backup hosts. These servers are typically 20% of the total number of servers in the Cassandra cluster. These servers are used to deduplicate the data from Cassandra cluster during backup and optimize the backup process. During a backup or restore, Cassandra keyspace are streamed in-parallel between the Cassandra cluster and the data staging servers. The data staging servers, represent a staging cluster. You need to deploy the nodes wherein, they are used depending on the size of data that needs to be backed up or restored.
Data reduction	As part of data reduction the following tasks are performed: ■ Efficient reconciliation Efficient reconciliation data for same keys from different nodes are transferred to the same node in the backup nodes. Reconciliations happen in-parallel within each data staging servers without any inter-node communication. ■ Record synthesis While iterating over the records, columns of the same key from different SStables are merged. ■ Semantic Deduplication Stale and duplicate records (replicas) are identified and removed.
NetBackup primary server	All the jobs are executed from the NetBackup primary server.
Parallel streams	The NetBackup parallel streaming framework allows data blocks from multiple nodes to be backed up using multiple backup hosts simultaneously.

Upgrading Cassandra from NetBackup 10.0

Cassandra support is added from NetBackup 10.0 with CLI support only and with policy-based backup and restore. NetBackup 10.1 and later releases have Cassandra support using NetBackup APIs and Web UI. It is based on protection plan methodology.

When upgrading from NetBackup 10.0, if you have policies for Cassandra backups, perform the following:

- You must delete these policies. Stop taking backups using the CLI as they are not supported.
- Configure the DSS and Apache Cassandra cluster from the Web UI and start with a fresh backup schedule.
- Ensure that the primary, media, clients are on NetBackup 10.1 or later. Back-level media or clients are not supported for Cassandra in NetBackup 10.1 and later.

When you restore old backup images of NetBackup 10.0, ensure that you have the Cassandra clusters and staging cluster configured in the NetBackup Web UI.

The restore of the backup images created by NetBackup 10.0 will need to be restored using the NetBackup API interface for Cassandra recovery. Refer to the API details for the payload of the recovery APIs.

- Ensure that you have the credentials of the target Cassandra cluster. Refer to NetBackup 10.0 admin guide for details. Adding Cassandra credentials in NetBackup
- Ensure that the Pre-requisites for Cassandra Restore are met.
- Perform the configuration for Cassandra if the target Cassandra was not already configured. Refer to Configurations for Cassandra Restore.
- Ensure that specify the correct recover pay load for selection and rename. For rename refer to the supported Restore combinations listed.

Chapter 2

Managing Apache Cassandra

This chapter includes the following topics:

Add DSS Clusters

During a backup or restore, Cassandra key space are streamed in-parallel between the Cassandra cluster and the DSS cluster. Follow the procedure to add DSS cluster.

1. On the left pane, click **Apache Cassandra**.
2. Select **DSS Cluster** tab.
3. Click **Add** to add a DSS cluster.

   Note: A prerequisite window appears to add cluster with a downloadable template.

4. Click **Start**.

5 On the **Basic Properties** tab, enter the following:

- **DSS Cluster name**
  The DSS cluster name must follow the limit of 256 characters.
- **CBR node host name**
- **CBR node key**

Note: To obtain this node key run the `cat /etc/ssh/ssh_host_rsa_key.pub |awk '{print $2}' |base64 -d |sha256sum |awk '{print $1}'` command. This node key must contain 64 characters.

---

6 Click **Next.**

7 On the **DSS cluster nodes** tab, add one or more cluster nodes. You can add multiple nodes either in chip format or comma separated.

Note: In case of duplicate entry, only one is entry is considered.

---

8 Click **Next**.

9 On the **Credentials** tab, do one of the following:

Select existing credentials:

- Search the desired credentials and select from the list.

Add new credentials:

- Select **Add new credential** and enter the following details.
  - Credential name
  - Tag
  - Description
  - Host username
  - Host password
  - Database username
  - Database password
  - Authentication type

Note: For **SSL** authentication, JMX credentials are mandatory.

Note: Credential name must follow the limit of 256 characters. Tag and Description are optional.

- Click **Next**.
- Click **Add**.
- On the **Credential Permission** tab, select a role to provide permissions for credential.
- Select the Permission from the following options.
  Permissions varies per selected role.
  - View
  - Create
  - Update
  - Delete
  - Manage Access
  - Assign Credentials
- Click **Save**.

10 Click **Next**.

11 On the **Backup hosts** tab, from **Primary backup host**, search and select the host.

Note: Any RHEL media server or RHEL client can be used as the backup host.

12 To add additional backup hosts, click **Add,** and select one or more host.

Note: You can also use NetBackup client as a backup host.

13 Click **Next**.

14 On the **Setting** tab, select the following:

- **DSS distribution**
  Thin-client distribution directory on the data staging servers. The path must be in UNIX format.
- **Script home**

The value is used for CBR package installation on the Apache Cassandra nodes.

- **Working directory**
  The folder where the thin client would stage the data and process them.

Note: Ensure that all the paths configured has read and write access for the credentials specified in the DSS cluster and Cassandra cluster.

---

15 On the **Advanced setting** page, review and make the necessary changes in the following:

- **Job cleanup time out**
  The time-out to the typical time it takes to back up the cluster.
- **DSS minimum RAM**
  The minimum RAM requirement for Data optimization on data stage server.
- **DSS minimum storage per backup node**
  The minimum storage requirement for Data optimization on data stage server.
- **Concurrent compaction**
  The maximum number of compactions that can run concurrently.
- **Loader memory size**
  The heap memory size for Cassandra table loader.
- **Concurrent transfer**
  The value is used to transfer parallel data from production to the data stage server. Default value is 8.

16 Click **Next**.

17 Review the data and click **Add**.

Edit DSS Clusters

Use this procedure to make any changes to the existing DSS cluster's configuration.

1 On the left pane, click **Apache Cassandra**.

2 Select **DSS Cluster** tab.

3 Select the desired cluster.

4 On the right corner of the screen, select **Edit setting** and make the required changes in the Path settings and/or Advanced settings.

5 Select **DSS Nodes** tab and select the required inline actions.

The following are the inline actions available.

Export as CSV:

- Use **Show or hide columns** to select the required fields.
- Select **Export as CSV**.
  A file is downloaded with the filtered details.

Edit CBR node:

- Select **Edit CBR node.**
- Select new CBR node from the list.

6 Click **Save**.

Remove DSS Clusters

Use this procedure to remove the existing DSS clusters.

1 On the left pane, click **Apache Cassandra**.

2 Select the **DSS Clusters** tab.

3 Select the cluster to remove.

4 Click **Remove**.

Add Apache Cassandra Clusters

Use this procedure to add Apache Cassandra cluster.

1 On the left pane, click **Apache Cassandra**.

2 Select **Apache Cassandra Cluster** tab.

Note: Prerequisite pop-up to add cluster appears.

3 Click **Start.**

4 On the **Basic properties** tab, do the following:

- Enter **Apache Cassandra Cluster name**.

Note: The name must not have any white space, special characters, or non-English characters.

- Select **DSS Cluster**.
  Used as data staging servers to protect the specified cassandra cluster.
- Enter **Discovery node host name**
  Used to discover all the nodes in the Apache Cassandra cluster.
- Enter **Discovery node key.**

Note: To obtain the RSA key run the `cat /etc/ssh/ssh_host_rsa_key.pub |awk '{print $2}' |base64 -d |sha256sum |awk '{print $1}'` command. This node key must contain 64 characters.

---

5 Click **Next.**

6 On the **Credentials**, do one of the following:

Select the credentials from the list:

- Search and select the required credentials from the available list.

Add new credential:

- Select Add new credential and enter the following details.
  - Credential name
  - Tag
  - Description
  - Host username
  - Host password
  - Database username
  - Database password
  - Authentication type

Note: For **SSL** authentication, JMX credentials are mandatory.

Note: Credential name must follow the limit of 256 characters. Tag and Description are optional.

- Click **Next**.
- Click **Add**.

- On the **Credential Permission** tab, select a role to provide permissions for credential.
- Select the permission from the following options.
  Permissions varies per selected role.
  - View
  - Create
  - Update
  - Delete
  - Manage Access
  - Assign Credentials
- Click **Save**.

7 Click **Next.**

8 On the **Setting** tab, enter the following:

- **Backup datacenter name**
  Used to protect the cluster.
- **Node down threshold**

9 Click **Next**.

10 Review the data and click **Add**.

Edit Apache Cassandra Clusters

1 On the left pane, click **Apache Cassandra.**

2 Select the **Apache Cassandra Cluster** tab.

3 Select the cluster to edit.

4 On the right corner of the screen, click **Edit Cluster**.

5 Review and make the necessary changes.

6 Click **Save**.

Remove Apache Cassandra Clusters

1 On the left pane, click **Apache Cassandra.**

2 Select the **Cassandra Cluster** tab.

3 Select the required Cluster.

4 Click **Remove**.

Chapter 3

Protect

This chapter includes the following topics:

Protecting Apache Cassandra assets

To protect the Apache Cassandra cluster on a very high level, you need:

- NetBackup primary server.
- NetBackup media server.
- A backup host that is NetBackup primary, NetBackup media server, or a NetBackup client.

For the supported primary and media server configurations, refer the NetBackup compatibility list.

The backup host that is NetBackup media server or a NetBackup client for Cassandra is supported only on an RHEL.

NetBackup Appliance, NetBackup Flex Appliance and NetBackup Flex Scale are supported as a NetBackup primary, media server, or as a client that can act as a backup host.

Use the following procedure to protect Apache Cassandra assets.

1. On the left pane, from **Protection**, select **Protection Plans**.
2. Click **+ Add** to add protection plan.
3. On the **Basic properties** tab, enter the following:
   - **Name**
   - **Description**
4. From the workload, select **Apache Cassandra**.

5 Click **Next**.

6 On the **Schedule** tab, click **Add Schedule** and select the following attributes:

- **Backup type**.

  Note: Select either Full or Differential incremental. Full backup backs up all the files specified in the selection list. Where as differential backup backs up the specified file changes.

- **Recurrence**
- **Keep for**

  Note: Select either Replicate the backup or Duplicate a copy immediately to long-term retention.

Note: To add another attribute, click Add and add another or click Add.

7 Move to **Start Window.**

8 Select the date and time to start the backup.

Note: Default, Backup storage is none.

9 To select a storage, click **Edit**.

10 Select the **storage** and click **Use selected storage**.

11 Click **Next**.

12 On the **Permissions** tab, click **Add** and select the following:

- **Role.**
- **Permission.**

  Following are the available options:

  - View
  - Create
  - Update
  - Delete

- Manage Access
- Edit attributes
- Edit full and incremental schedules
- Edit transaction log schedules
- Subscribe

13. Click **Save**.
14. To modify the values, click **Edit** and make the required changes.
15. Review the data and click **Finish**.

Chapter 4

Pre-recovery Check

This chapter includes the following topics:

About the Pre-recovery Check

The pre-recovery check verifies the following:

- Backup host, data staging-server, Cassandra clusters is on RHEL.
- Availability of a DSS and Cassandra Clusters with the same display name.
- Availability of free space on all the Data staging servers in the DSS cluster.
- Cassandra service is up and running on all the data staging servers.
- DSS and Cassandra Cluster's credential validation.

Chapter 5

Recover

This chapter includes the following topics:

Recovering Apache Cassandra assets

1. On the left pane, select **Apache Cassandra**.
2. Locate and click **Cassandra Cluster**.
3. Click the **Recovery points** tab.

Note: You can recover a Cassandra cluster either to an original backup cluster or to a different cluster.

Restore cluster

1. From the Actions menu, select **Restore cluster**.
2. On **Recovery target** tab, do one of the following:
   - To restore the cluster to the original cluster, select **Restore to the original cluster**.
   - To restore the cluster to the different cluster, select **Restore to a different cluster** and select the target cluster from the list.
3. Click **Next**.
4. On **Recovery options** tab, to restore the key spaces, do one of the following:
   - Select **Restore keyspaces with original attributes** to continue with the same key spaces.

- Select **Edit keyspaces** to edit.
  - Click **Add keyspaces** to edit.
  - Select the desired key spaces and enter the following:
    - **New key space name**
    - **Strategy name**
    - **Datacenter name**

      Note: If you select the option Simple in the Strategy name, Datacenter name is automatically selected.

    - **Replica**
      Specify the number of replicas that you required for the key space.

5 Select **Next**.

6 On **Review** tab, review the values and if required make changes to the Recovery target and Recovery options values.

7 Click **Start recovery**.

Restore keyspaces and column families.

1 From the Actions menu, select **Restore keyspaces and column families**.

2 Click **Add**.

Add key spaces and column families window displays.

3 Select the required Keyspace and family and Click **Select**.

4 Click **Next**.

Note: Backup and restore of system keyspaces, and respective column families are not supported . Cassandra has user defined data types like Materialized views are not being restored while performing restore operation.

SSH key fingerprint

Use `cat /etc/ssh/ssh_host_rsa_key.pub |awk '{print $2}' |base64 -d |sha256sum | awk '{print $1}'` command on target host to get the SHA256-based RSA key.

Note: Similarly, change the public key path, run the command to get ecdsa or DSS SSH key fingerprint configured on target host.

Recover from a copy of the recovery point

Use this procedure, if you have one or more copies of the recovery points of Cassandra to recover from those copies.

To select a copy for restore.

1. On the left, click **Catalog**
2. Select the correct required details such as date and time range and client.
3. Search for backup image using **Copies** for example, **Copy2** or **Copy3**.
4. Select the desired image and click the **Set primary copy** option next to the image.

   Note: Once this copy becomes a primary copy, it is used for restore.

Chapter 6

Troubleshoot

This chapter includes the following topics:

Troubleshooting Apache Cassandra issues

For more information about Apache Cassandra troubleshooting, check the following details:

- Check the **Job details** section of the job in **Activity monitor** for failures.
- Check the following logs:
  - `bprd`
  - `bpVMutil`
  - `nbaapireq_handler`
  - `bpbkar`
  - `tar`
  - `nbaapidiscv`

Errors and recommended actions

The following table describes the problem that might occur.

Table 6-1 Error and recommended actions

Error message or cause	Explanation and recommended action
CBR node not reachable.	The CBR node in the DSS cluster is not reachable. Recommended actions: ■ Verify the node that specified as the CBR node is up, and reachable from the backup host. ■ If required, add a new node to the DSS cluster to change the CBR node.
CBR RSA key mismatch.	The CBR node RSA key specified do not match with the RSA key provided by the CBR node. Recommended actions: ■ Do one of the following to specify the correct RSA key for the CBR node. ■ Add a new node in the DSS cluster. ■ Select the CBR node and enter the respective RSA key.
CBR is reachable, but some nodes not reachable.	Some of the nodes in the DSS cluster are not reachable. Recommended actions: ■ Ensure that the nodes which are mentioned in the DSS cluster are up and running and are reachable from the backup host. ■ Add new nodes from the DSS cluster details.
Invalid nodes in the cluster. Credentials are invalid. Apache Cassandra not installed.	The nodes in the clusters are invalid or unable be log in indicates the possibility that Apache Cassandra is not installed. Recommended actions: ■ Ensure that correct credentials are entered on clusters. ■ Select a different account for the cluster credentials in NetBackup. **Note:** The credentials must be valid on all the cluster nodes. ■ Ensure that Apache Cassandra is installed on all the nodes. **Note:** All nodes must have the same version of Apache Cassandra.

Table 6-1 Error and recommended actions *(continued)*

Error message or cause	Explanation and recommended action
CBR credentials mismatch.	The DSS cluster's credentials are invalid on the CBR node. Recommended actions: ■ Ensure that correct credentials are entered on the CBR node. ■ Select a different credential for the DSS cluster in NetBackup. **Note:** The credentials must be valid on all the cluster nodes.
DSS node credentials are invalid.	The credentials that specified for the DSS node are not valid. Recommended actions: ■ Ensure that correct credentials are entered on the DSS node. ■ Select a different account for the DSS cluster in NetBackup. **Note:** The credentials must be valid on all the cluster nodes.
Discovery nodes not reachable.	The Discovery node for the Cassandra cluster is not reachable. Recommended actions: ■ Ensure that the specified node is up and running. ■ Ensure that the Apache Cassandra service is up and running on the node. ■ Edit the Apache Cassandra cluster settings to change the discovery node.
Discovery node RSA mismatch.	The status of the cluster is **Pending for approval.** Recommended action: Approve the RSA key of discovery node.

Table 6-1 Error and recommended actions *(continued)*

Error message or cause	Explanation and recommended action
Discovery node credentials invalid.	The credentials of the Cassandra cluster are invalid for the discovery node. Recommended actions: ■ Ensure that correct credentials are entered on the discovery node. ■ Select a different credential for the Cassandra cluster in NetBackup. **Note:** The credentials must be valid on all the nodes in the Cassandra cluster.
Apache Cassandra not installed on the discovery node.	The Apache Cassandra is not installed on the discovery node of the Cassandra configuration. Recommended actions: ■ Select a different node for discovery in the Apache Cassandra cluster.
Apache Cassandra cluster name not matched.	The discovery node belongs to a different cluster. Recommended actions: ■ Enter the correct cluster name in Cassandra cluster configuration in NetBackup. ■ Select the correct discovery node for the cluster.
Datacenter name not matched.	The data center name of the Cassandra nodes mismatch. Recommended action: ■ Change the name of the data center in Cassandra cluster configuration in NetBackup.
Cassandra Cluster pending approval.	One ore more nodes in the Cassandra cluster and RSA key require approval. Recommended action: ■ Review and approve the RSA key on the node.

Table 6-1 Error and recommended actions *(continued)*

Error message or cause	Explanation and recommended action
Cassandra nodes invalid credentials.	Cassandra cluster credentials mismatch with one or more nodes in the cluster. Recommended action: ■ Select the correct credentials for the Cassandra cluster. **Note:** The credentials must be valid on all the Cassandra cluster.
The Cassandra DSS cluster and/or Cassandra production cluster goes into invalid status due to concurrent operations.	NetBackup can do only one operation such as discovery, validate, backup, restore, at any given point in time on the DSS and production clusters. If one operation is running, the other fails and put the cluster in invalid state. If the cluster is in invalid state while the previous operation is running, wait for the pervious operation to complete. Then retry the operation you previous operation like discovery, validation, backup now or restore.
Discovery of more than one cluster take time and doesn't show the discovering status immediately.	NetBackup works, discovery one cluster at a time hence the state of the second cluster changes after the first one is complete and reflect the appropriate status after completion.
Cassandra restore page becomes unresponsive for keyspaces/column families more than 300.	Use computer with higher configuration for example RAM size is 16GB or more.
Incremental backup fails with Error 6.	Recommended action: Verify if the scheduled full backup is completed. **Note:** Full backup run with a use of **Backup now** is not considered as it is not linked to scheduled backup.
Backup/restore does not work post primary cluster node failover.	Recommended action: For primary server cluster, after the fail-over of node, update the primary server's `bp.conf` with `APP_PROXY_SERVER = NetBackup media server`. **Note:** NetBackup media server is used as the backup host.

Table 6-1 Error and recommended actions *(continued)*

Error message or cause	Explanation and recommended action
no java in (/sbin:/bin:/usr/sbin:/usr/bin) Java executable not found (hint: set JAVA_HOME)	Recommended action: Update the `secure_path` list with Java executable path in `/etc/sudoers` file.

Chapter 7

API for Cassandra

This chapter includes the following topics:

Using APIs to manage, protect, or recover Cassandra assets

In this document, you will find details about NetBackup APIs for the Cassandra workload operations. The operations include asset management, filtering, sorting, recovery management and protection plan management. For each API end point, the user can find detailed information about the requests, responses, errors, and payloads in this document.

See the following for information on the APIs:

- All the NetBackup APIs are listed at the following location:
  Services and Operations Readiness Tools (SORT) > Knowledge Base > Documents

Create or update Cassandra assets

The API creates the requested Cassandra assets - Cassandra cluster, Cassandra node, DSS cluster, and DSS node.

The request payload can contain detailed information about a single asset or a combination of assets. A Cassandra asset can be created or updated in two steps:

- To retrieve the generated ID, a POST request must be made.
- It is necessary to make a GET request, which saves the asset in the database.

Table 7-1 1. Request parameters for Cassandra Cluster and Cassandra Nodes

API	Important variables and options

`https://PrimaryerverName`
 	
`/netbackup/asset-service/queries/`
	■ `Workloads`: cassandra
	■ `assetType`: Defines the asset type - cluster
	■ `credentialName`: Credential name used for asset service.
	■ `workloadType`: cassandra
	■ `clusterName`: Valid Cassandra Cluster name.

`https://PrimaryServerName` `/netbackup/asset-service/queries/`	■ `Workloads`: cassandra ■ `assetType`: Defines the asset type - node ■ `workloadType`: cassandra

Table 7-2 2. 3. Request parameters for Data Staging Cluster and Nodes

API	Important variables and options
`https://PrimaryServerName` `/netbackup/asset-service/queries/`	■ `Workloads`: cassandra ■ `assetType`: Defines the asset type - cluster, node, dataStagingCluster ■ `workloadType`: cassandra ■ `clusterName`: Valid Cassandra Cluster name.
`https://PrimaryServerName` `/netbackup/asset-service/queries/`	■ `Workloads`: cassandra ■ `assetType`: Defines the asset type - dataStagingNode. ■ `workloadType`: cassandra

Table 7-3 Response parameters

API	Important variables and options

`https://PrimaryServerName`
`/netbackup/asset-service/queries/`
`Response parameters`
	■ `type` Name of the query response type.
	■ `id`: Response Id for asset creations/modification
	■ `links`: Complete asset link

Method: Get

Assets are stored using this API in the NetBackup database. It uses the GET method with a valid asset response ID in the URL.

Table 7-4 Response of storing Assets to Database

API	Important variables and options
`https://PrimaryServerName /netbackup/asset-service/queries/ Asset Response ID`	■ `id`: Response Asset ID for create-or-update-assets query ■ `status`: Asset Response status - Success, Failed and In progress ■ `percentComplete`: Percent of completion ■ `totalWorkItems`: Number of Response Items ■ `totalSuccessfulWorkItems`: Number of successful items ■ `totalFailedWorkItems`: Number of failed items ■ `totalInProgressWorkItems`

Delete Asset Payload

A specific asset or all assets can be deleted from the NetBackup database using this API. An individual **assetType**, such as a node or **dataStagingNode**, can be deleted. It deletes all corresponding child nodes associated with a cluster or **dataStagingCluster** if you use this API to delete **assetType** as cluster or **dataStagingCluster**.

Table 7-5 Request Payload for Deleting Assets

API options	Important variables and options
`https://PrimaryServerName /netbackup/asset-service/queries`	■ `type` Name of the query response type. ■ `queryName` Name of the query to create or update different assets ■ `workloads` - cassandra ■ `correlationId` - It is an identifier for tracking a work item ■ `id` - Id assigned to the asset ■ `assetType` - Type of Asset {node, dataStagingNode, dataStagingCluster, cluster}

Request Payload for removing Assets from Database

Database assets can be deleted using this API. This is the 2nd request used with the delete response id received from the previous request.

Table 7-6

Method	URL
GET	`https://PrimaryServerName /netbackup/asset-service/queries/ Asset Delete Response ID` For example: *bcc0eb1f-6613-427d-8101-19376fd689f7*

Table 7-7 Response of removing Assets from Database

API	Important variables and options
`https://PrimaryServerName /netbackup/asset-service/queries/ Asset Delete Response ID` For example: *bcc0eb1f-6613-427d-8101-19376fd689f7*	■ `type` Name of the Operation response type ■ `id` - Response Id for Generic delete-assets ■ `status` - - Asset Response status - SUCCESS ■ `percentComplete` - Percent of completion ■ `totalWorkItems` - No of total Response work Items ■ `totalSuccessfulWorkItems` - No of total Success work Items ■ `totalFailedWorkItems` - No of total failed work Items ■ `totalInProgressWorkItems` - No of total in progress work Items ■ `correlationId` - It is an identifier for tracking a work item ■ `status` - Asset response status ■ `message` - Asset response message ■ `percentComplete` - Percent of completion ■ `offset` ■ `limit` ■ `hasNext` ■ `first`

Table 7-8 Request for Get all Asset

`https://PrimaryServerName /netbackup/asset-service/ workloads/cassandra/assets`	`Assets`

Response of Get all Assets

All the assets are listed on default descending order, refer Get asset by ID table for response parameter details. According to the Get asset by ID table, all assets are listed in default descending order.

Get asset by ID

This API is used to get specific asset details. Details about specific assets can be obtained using this API.

Method: GET

Table 7-9

API	Important variables and options
`https://PrimaryServerName/ netbackup/asset-service/workloads/ cassandra/assets/Asset ID`	■ `type` - Name of the Operation response type ■ `id` - Response Id for Generic delete-assets ■ `assetType` - Defines type of asset ■ `displayName` - Name of asset ■ `providerGeneratedId` - Auto generated ID for asset ■
`https://PrimaryServerName/ netbackup/asset-service/workloads/ cassandra//Asset ID`	■ `parentProviderGeneratedId` - Auto generated id for nodes and dataStagingNodes ■ `clusterName` - cassndra

Filters

Filters can be applied on any of the below API endpoints:

- `https://PrimaryServerName/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster'`

- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'node'
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'dataStagingCluster'
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'dataStagingNode'
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'node' and parentProviderGeneratedId eq 'CASSANDRA_cluster_Cassandra10'
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster' and commonAssetAttributes/providerGeneratedId eq 'CASSANDRA_cluster_Cassandra1'
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster' and dataStagingClusterName eq 'DSS1'
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster' and backupDataCenterName eq 'earthCenter'
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'node' and nodeIp eq '10.221.110.234'
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=(assetType eq 'cluster') and (contains(tolower(clusterName), 'cassandra1'))
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=(assetType eq 'cluster') and (clusterName eq 'Cassandra1')&meta=accessControlId
- https://*PrimaryServerName*/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster' and clusterName eq 'Cassandra10'

Table 7-10

Filter Parameters	Filter Criteria	Value	Description
assetType/ queryName/ workloads/ correlationId/ credentialName/ workloadType/ displayName/ clusterName/ dataStaging ClusterName/ backupData CenterName/ nodeDownThreshold Percentage/ parentProvider Generated	eq	cluster, node, dataStagingCluster and dataStagingNode	Asset filter based on asset types
	eq	'node' and parentProviderGeneratedId eq 'CASSANDRA_cluster_Cassandra1'	
	eq	'cluster' and commonAssetAttributes/ providerGeneratedId eq 'CASSANDRA_cluster_Cassandra1'	
	eq	'cluster' and dataStagingClusterName eq 'DSS1'	
	eq	'cluster' and backupDataCenterName eq 'earthCenter'	
	eq	'node' and nodeIp eq '10.221.105.94'	
	eq	(assetType eq 'cluster') and (contains(tolower(clusterName), 'cassandra1'))	
	eq	(assetType eq 'cluster') and (clusterName eq 'Cassandra1')&meta=accessControlId	

Sorting

Sorting can be applied on any of the below API endpoints:

- `https://`*`PrimaryServerName`*`/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster'&sort=clusterName`

- `https://PrimaryServerName/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster'&sort=-clusterName`
- `https://PrimaryServerName/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster'&sort=commonAssetAttributes.credentials.credentialName`
- `https://PrimaryServerName/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'cluster'&sort=-commonAssetAttributes.credentials.credentialName`
- `https://PrimaryServerName/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'node'&sort=parentProviderGeneratedId`
- `https://PrimaryServerName/netbackup/asset-service/workloads/cassandra/assets?filter=assetType eq 'node'&sort=-parentProviderGeneratedId`

Table 7-11

Filter Parameters	Filter Criteria	Value	Description
assetType	eq	cluster'&sort=clusterName	
assetType	eq	node and parentProviderGeneratedId eq 'CASSANDRA_cluster_Cassandra1	
assetType	eq	cluster'&sort= commonAssetAttributes .credentials.credentialName	
assetType	eq	'cluster'&sort=-commonAssetAttributes. credentials.credentialName	
assetType	eq	'node'&sort=	
assetType	eq	parentProvider GeneratedId	

Table 7-11 *(continued)*

Filter Parameters	Filter Criteria	Value	Description
assetType	eq	'node' &sort=-parentProvider GeneratedId	

Protection plan

The below APIs create or update the policy which is used to trigger actual backup operation on the associated asset. These APIs help to create or update the policy for both incremental and full backup. Basis the created policies, the backup jobs are automatically triggered based on the schedule details provided.

Table 7-12 Create SLO : Full and Differential Incremental

Method	Post
`https://PrimaryServerName/netbackup/servicecatalog/slos`	

Table 7-13 Request Payload for Creating SLO:

API	Important variables and options
`https://PrimaryServerName/netbackup/servicecatalog/slos`	▪ `name` ▪ `scheduleType` ▪ `backupStorageUnit` ▪ `dayOfWeek` ▪ `startSeconds` ▪ `durationSeconds` ▪ `frquencySeconds` ▪ `workloadType` ▪ `policyNamePrefix` ▪ `policyDefinition` ▪

Table 7-14 Response of Creating SLO

API	Important variables and options
`https://PrimaryServerName/netbackup/servicecatalog/slos`	■ `201` - Successfully created the SLO definition. ■ `400` - Bad request ■ `401` - The Authorization header is missing, the token is invalid, or you do not have permission for this action. ■ `409` - An SLO with the same name already exists. ■ `500` - An unexpected system error occurred.

Create Subscription

Table 7-15

Method	Post
`https://{{hostname}}/netbackup/servicecatalog/slos/{{newsloid}}/subscriptions`	

Table 7-16 Request Payload for Creating Subscription:

API	Important variables and options
`https://{{hostname}}/netbackup/servicecatalog/slos/{{newsloid}}/subscriptions`	■ `selectionId` ■ `selectionType`

Table 7-17 Response of Creating Subscription

API	Important variables and options
`https://{{hostname}}/netbackup/servicecatalog/slos/{{newsloid}}/subscriptions`	■ `201` - Successfully created the Subscription ■ `400` - Bad request ■ `401` - The Authorization header is missing, the token is invalid, or you do not have permission for this action. ■ `409` - An SLO with the same name already exists. ■ `500` - An unexpected system error occurred.

Discovery

The discovery APIs cater actual initiation or stopping of the discovery over given asset.

Table 7-18 Start discovery

API	Important variables and options
`https://{{HOSTNAME}}/netbackup/admin/discovery/workloads/cassandra/start`	■ `Data` ■ `Type` ■ `Attribute` ■ `serverName`

Table 7-19 Stop discovery

API	Important variables and options
`https://{{HOSTNAME}}/netbackup/admin/discovery/workloads/cassandra/stop`	■ `Data` ■ `Type` ■ `Attribute` ■ `serverName`

Recovery request

The Recovery APIs cater actual recovery of specific Cassandra asset, based on the details of recovery point, source, destination etc.

Table 7-20

API	Important variables and options
POST /recovery/workloads/cassandra/scenarios/cluster/recover	■ Data ■ Type ■ Attribute ■ recoveryPoint ■ client ■ filter ■ backupId ■ recoveryObject ■ clusterNewName ■ recoveryOptions ■ backupHost ■ additionalBackupHosts ■ nbu-backup-host2 ■ nbu-backup-host3 ■ nbu-backup-host4 ■ overwrite ■ restoreSelections ■ selectionType ■ selectionCriteria ■ keyspace

Table 7-21 Response of Recover API

API	Important Response codes
`POST /recovery/workloads/cassandra/scenarios/cluster/recover`	■ `201` - The recovery job started successfully. ■ `400` - Some mandatory attributes were not found or the specified client, backup image, or the input JSON was invalid. ■ `401` - The Authorization header is missing, the token is invalid, or you do not have permission for this action. ■ `404` - The specified client or backup image was not found ■ `406` - Invalid Accept type. Make sure your Accept header matches what this API produces ■ `415` - Unsupported Media Type. The media type specified in the Content-Type header is not supported by this API. ■ `500` - Internal server error. Failed to start recovery. ■ `503` - The server is busy. Failed to start recovery.