NetBackup™ for Greenplum Administrator's Guide

Release 11.2

VERITAS
now part of Cohesity

NetBackup™ for Greenplum Administrator's Guide

Last updated: 2026-05-28

Legal Notice

Cohesity Support

Reach Cohesity Support

There are several ways to create a Cohesity support case.

- Go to Cohesity Support, to search in our knowledge base; or contact us by phone - United States and Canada: 1-855-9CO-HESI (926-4374), option 2.
- Log in to the Cohesity Support Portal to create a new case.
- Click the (**?**) icon on the Cohesity UI and select Support Portal.

Support/Service Assistance

First, contact the Service Provider that you have contracted for service and support. If you work directly with Cohesity and have a product warranty/entitlement, repair pricing, or technical support-related question, see your options below:

- To find solutions to your product issues or for suggestions or best practices, visit the Cohesity Knowledge Base.
- Log in to the Cohesity Support Portal to create a new case.
- To monitor your open cases, log in to the portal and click the **Cases** tab on the home page. This page should have all the case statuses and updates. You can also view individual case status.

Cohesity Software Running on Partner Hardware

For Cohesity software running on qualified third-party hardware, the following support workflow applies:

1. The customer may contact Cohesity Support first if the issue cannot be determined as a hardware issue.

   Note: Cohesity cannot process hardware replacement requests for partner hardware.

2. Cohesity Support triages the issue. If it is a software issue, Cohesity Support continues to work on it.
3. If it is a hardware/firmware issue or is suspected to be a hardware/firmware issue, Cohesity provides information about the issue to the customer and requests that the customer open a support ticket with the appropriate partner.
4. If needed, Cohesity Support can join a three-way call with the partner and the customer.
5. The customer informs Cohesity Support on the progress of the partner's case.

Contents

Chapter 1

Introducing NetBackup for Greenplum

This chapter includes the following topics:

About NetBackup for Greenplum

The NetBackup for Greenplum plug-in integrates the database backup and recovery capabilities of NetBackup for Greenplum with the backup and the recovery management capabilities of NetBackup. The plug-in uses the Greenplum utilities `gpbackup` and `gprestore`.

In this guide, the Greenplum Database is referred to as Greenplum. NetBackup for Greenplum Database is referred to as NetBackup for Greenplum.

See “Features of NetBackup for Greenplum” on page 7.

NetBackup for Greenplum terminology

Table 1-1 shows the terms that might be new to a Greenplum database administrator or a NetBackup administrator.

Table 1-1 Terminology for Greenplum

Term	Meaning
Coordinator	The entry to the Greenplum Database system. It accepts client connections and SQL queries and distributes work to the segment instances. [1]
Segments	The independent PostgreSQL databases that each store a portion of the data and perform the majority of the query processing. [2]
Segment_hosts	The host that contain the segments.
Greenplum `gpbackup` and `gprestore` utilities	These Greenplum utilties that are used for Greenplum database backups and database restores. The `gpbackup` command is used to back up. The `gprestore` command is used to restore.
Greenplum backup scripts	The shell scripts that control the NetBackup for Greenplum operations.

[1] About the Greenplum Architecture

[2] About the Greenplum Segments

Features of NetBackup for Greenplum

Table 1-2 shows NetBackup for Greenplum's main features and introduces some terms that are used in this documentation.

Table 1-2 NetBackup for Greenplum features and descriptions

Feature	Description
Media and device management	All the devices Media Manager supports are available to NetBackup for Greenplum.
Scheduling facilities	NetBackup scheduling facilities on the primary server can be used to schedule automatic and unattended Greenplum backups. This feature also lets you choose the times when these operations can occur. For example, to prevent interference with normal daytime operations, you can schedule your database backups to occur only at night.
Multiplexed backups and restores	NetBackup for Greenplum lets you take advantage of NetBackup's multiplexing capabilities. Multiplexing directs multiple data streams to one backup device, thereby reducing the time necessary to complete the operation.

Table 1-2 NetBackup for Greenplum features and descriptions *(continued)*

Feature	Description
Transparent Greenplum and regular file system backup and restore operations	All backups and restores run simultaneously and transparently without any action from the NetBackup administrator. The database administrator can run database backup and restore operations through NetBackup. An administrator or any other authorized user can use NetBackup to run database backups and restores.
Sharing the same storage units that are used for other file backups	Devices and media can be shared for other backups or to you can give Greenplum exclusive use of certain devices and media. NetBackup for Greenplum can use Media Manager, disk, and Media Server Deduplication Pool (MSDP) storage units.
Centralized and networked backup operations	From the NetBackup primary server, you can schedule database backups or start them manually for any client. The Greenplum databases can also reside on any hosts that are different from the devices on which NetBackup stores the backups.
User interfaces	NetBackup provides the NetBackup web UI for monitoring Datastore policy jobs in the Activity monitor and for policy management.
Parallel backup and restore operations	NetBackup for Greenplum supports the parallel backup and restore capabilities of Greenplum. For example, a user can run more than one tape device at a time for a single Greenplum backup or restore. This usage can reduce the time necessary to complete the operation.

NetBackup for Greenplum technical overview

NetBackup for Greenplum uses the `gpbackup` and `gprestore` commands to backup and restore Greenplum databases. These commands are not part of the Tanzu Greenplum database software installation. The *Tanzu Greenplum Backup and Restore software* package contains the `gpbackup` and `gprestore` commands so it must be installed with the database software. The NetBackup for Greenplum agent is a command-line plug-in for `gpbackup` and `gprestore`. (The NetBackup for Greenplum plug-in command-line is `nbgpbackup`.)

Greenplum is a distributed database that runs on multiple hosts. The database uses one coordinator host and multiple segment hosts. Each segment host can contain multiple segments. Each host must have the NetBackup client software installed. The Greenplum agent uses the Datastore policy type.

For more information on the Greenplum Database, see your Greenplum documentation.

The `gpbackup` command calls `nbgpbackup` multiple times during a backup. On the segment hosts `nbpgbackup` is called many times to protect a database. For the best performance, the agent attempts to put multiple files from a segment into an image. To achieve this action the agent uses shared memory to pass the data to a `nbgpbackup` process that communicates with the NetBackup servers. Each NetBackup stream is a separate process that services its shared memory. After the segment host backups are complete `gpbackup` also calls `nbgpbackup` multiple times on the coordinator host. But unlike with the segment hosts, each of those calls results in separate backup images.

Each segment on each segment host is backed up into a separate image. To achieve this action the `number_of_streams` option in the `config` file must be at least equal to the number of segments per host. If more than one stream per segment is configured, then there can be multiple images for each segment.

Sequence of operation for NetBackup for Greenplum backups

Greenplum script files control the NetBackup operations. You use the NetBackup web UI to configure a schedule to use a script to perform NetBackup operations.

The following process takes place when a script is selected for a backup:

- A NetBackup process that is called `bphdb` starts the Greenplum backup script on the client.
- The Greenplum backup script starts the `gpbackup` utility.
- Greenplum starts the requested backup operation on the databases.
- `gpbackup` calls the `nbgpbackup` command line and starts passing its data to back up.
- `nbgpbackup` sends data to NetBackup for Greenplum, which transfers data to the media server.
- The media server sends the data to a storage unit.

A restore works in essentially the same manner except NetBackup uses `gprestore` instead of `gpbackup`. That command instructs the media server to retrieve the data from the storage unit and send it to NetBackup for Greenplum on the client.

Chapter 2

Installing NetBackup for Greenplum

This chapter includes the following topics:

Verifying the operating system and platform compatibility

Verify that the NetBackup for Greenplum agent is supported on your operating system or platform.

To verify operating system and compatibility

- Go to the NetBackup compatibility list site.

  https://support.cohesity.com/s/article/article-100040093

NetBackup server and client requirements

Before you install NetBackup, review the requirements for the NetBackup server and the NetBackup clients.

NetBackup server requirements

Verify that the following requirements are met for the NetBackup server:

- The NetBackup server software is installed and operational on the NetBackup server.

See the NetBackup Installation Guide.

- Make sure that you configure any backup media that the storage unit uses. The number of media volumes that are required depends on several things:
  - The devices that are used and the storage capacity of the media.
  - The sizes of the databases that you want to back up.
  - The amount of data that you want to archive.
  - The size of your backups.
  - The frequency of backups or archives.
  - The length of retention of the backup images.

  See the NetBackup Administrator's Guide, Volume I.

NetBackup client requirements

Verify that the following requirements are met for the NetBackup clients:

- The client software is installed on each node of the Greenplum configuration.
- The client software is installed on each computer that you want to back up.
- To use the new features that are included in NetBackup for Greenplum in NetBackup 11.2, you must upgrade your NetBackup for Greenplum clients to NetBackup 11.2. The NetBackup media server must use the same version as the NetBackup for Greenplum client or a higher version than the client.

Greenplum server software requirements

Verify the following regarding the Greenplum server software on the NetBackup server or client:

- The Greenplum server software must be installed and operational.
  Refer to the Application/Database Agent Compatibility List for supported versions of the Greenplum server software.
- VMware Tanzu Greenplum Backup and Restore package, version 1.30 or later. You can download this package from https://www.support.broadcom.com.

Chapter 3

Configuring NetBackup for Greenplum

This chapter includes the following topics:

About NetBackup for Greenplum configuration

Before you attempt to configure NetBackup for Greenplum, review the installation information *Installing NetBackup for Greenplum*.

After you complete the installation, you can follow the procedures in Table 3-1 to configure your environment.

Table 3-1 Steps to configure NetBackup for Greenplum

Step	Action	Description
Step 1	Configure a backup policy.	See "About configuring a backup policy for Greenplum " on page 13.
Step 2	Create the backup scripts (shell scripts).	See "About the backup script for NetBackup for Greenplum" on page 20.
Step 3	Review the best practices.	See "Best practices for Greenplum backups" on page 21. See "Use of the --jobs option or backup streams and the effect on NetBackup jobs and throttling options" on page 22.
Step 4	(UNIX and Linux) Configure the permissions for log files.	See "About permissions for NetBackup for Greenplum log files" on page 26.
Step 5	Test the configuration settings.	See "Perform a manual backup" on page 27.

About configuring a backup policy for Greenplum

A backup policy defines the backup criteria for a specific group of one or more clients.

These criteria include the following:

- Storage unit and media to use
- Policy attributes
- Backup schedules
- The clients to back up
- The script files to run on the clients

See "Add a policy for Greenplum" on page 13.

Add a policy for Greenplum

This topic describes how to create a policy to protect a database.

To add a policy for Greenplum

1. Open the NetBackup web UI.
2. On the left, select **Protection > Policies**. Then select **Add**.
3. Type a unique name for the new policy.

4 From the **Policy type** list, select **Datastore**.

5 Complete the entries on the **Attributes** tab.

See "About policy attributes" on page 14.

6 Add other policy information as follows:

- Add schedules.
  See "Configure an application backup schedule" on page 16.
  See "Configure automatic backup schedules" on page 17.
- Add clients.
  See "Add clients to a policy" on page 18.
- Add scripts to the backup selections list.
  See "Adding NetBackup for Greenplum scripts to the backup selections list" on page 19.

7 When you have completed the policy configuration, select **Create**.

About policy attributes

With a few exceptions, policy attributes for a Datastore policy are managed in the same way as for most other policy types. Certain policy attributes vary according to your specific backup strategy and system configuration.

For more information on policy attributes, see the NetBackup Administrator's Guide, Volume I.

Table 3-2 Policy attributes for NetBackup for Greenplum policies

Attribute	Description
Policy type	Determines the types of clients that can be backed up with the policy. For Greenplum databases, select the policy type **Datastore**.
Allow multiple data streams	This policy attribute is ignored for NetBackup for Greenplum. If you want to adjust the number of streams (or the jobs that are created), the number of streams should be equal to the number of segments. Refer to the `number_of_streams` option. See "The nbgpbackup_config.yaml file" on page 23.
Keyword phrase	For NetBackup for Greenplum, the **Keyword phrase** entry is ignored. You can set a keyword value in the `nbgpbackup_config.yaml` file.

Schedule properties

This topic describes the schedule properties that have a different meaning for database backups than for file system backups. Other schedule properties vary according to your specific backup strategy and system configuration. Additional information about other schedule properties is available. See the NetBackup Administrator's Guide, Volume I.

Table 3-3 Description of schedule properties

Property	Description
Type of backup	Specifies the type of backup that this schedule can control. The selection list shows only the backup types that apply to the policy you want to configure. See "NetBackup for Greenplum backup types" on page 16.
Schedule type	You can schedule an automatic backup in one of the following ways: ■ **Calendar** The **Calendar** option lets you schedule the backup operations that are based on specific dates, recurring week days, or recurring days of the month. ■ **Frequency** The **Frequency** specifies the period of time that can elapse until the next backup operation begins on this schedule. For example, assume that the frequency is 7 days and a successful backup occurs on Wednesday. The next full backup does not occur until the following Wednesday.
Retention	Specifies a retention period to keep backup copies of files before they are deleted. The retention level also denotes a schedule priority within the policy. A higher level has a higher priority. Set the time period to retain at least two full backups of your database. In this way, if one full backup is lost, you have another full backup to restore. For example, if your database is backed up once every Sunday morning, you should select a retention period of at least 2 weeks. The retention period for an application backup schedule refers to the length of time that NetBackup keeps backup images. The retention period for an automatic schedule controls how long NetBackup keeps records of when scheduled backups occurred. For example, if your database is backed up once every Sunday morning, you should select a retention period of at least 2 weeks.

Table 3-3 Description of schedule properties *(continued)*

Property	Description
Effect of type of schedule on retention period	The type of schedule you select affects the retention period as follows: ■ Frequency-based scheduling Set a retention period that is longer than the frequency setting for the schedule. For example, if the frequency setting is set to one week, set the retention period to be more than one week. The NetBackup scheduler compares the latest record of the automatic backup schedule to the frequency of that automatic backup schedule. This comparison is done to determine whether a backup is due. So if you set the retention period to expire the record too early, the scheduled backup frequency is unpredictable. However, if you set the retention period to be longer than necessary, the NetBackup catalog accumulates unnecessary records. ■ Calendar-based scheduling The retention period setting is not significant for calendar-based scheduling.

NetBackup for Greenplum backup types

Table 3-4 shows the backup types you that can specify for a NetBackup for Greenplum policy.

Table 3-4 NetBackup for Greenplum backup types

Backup type	Description
Application backup	The application backup schedule enables user-controlled NetBackup operations from the client. These operations include those initiated from the client and those initiated by an automatic schedule on the primary server. NetBackup uses the application backup schedule when the user starts a backup manually. Configure at least one application backup schedule for each policy. The Default-Application-Backup schedule is configured automatically as an application backup schedule.
Automatic backup	An automatic backup schedule specifies the dates and times for NetBackup to automatically start backups. NetBackup runs the scripts in the order that they appear in the file list. If there is more than one client in the policy, the scripts are run on each client.

Configure an application backup schedule

A database backup requires an application backup schedule. You cannot perform backups if this type of schedule is not included in the policy. NetBackup automatically creates this schedule and names it **Default-Application-Backup**.

The backup window for an application backup schedule must encompass the time period during which all scheduled jobs and client-initiated jobs can occur. This window is necessary because the application backup schedule accepts the backup request from NetBackup for Greenplum regardless of whether the backup was initiated from an automatic schedule or from the client. You can choose to set the window for the application backup schedule for 24 hours per day, seven days per week. This window ensures that your operations are never locked out due to the application backup schedule.

To configure an application backup schedule

1. Open the policy and select the **Schedules** tab.
2. Select the schedule that is named **Default-Application-Backup** and select **Edit**.
3. Specify the other properties for the schedule.

   See "Schedule properties" on page 15.
4. Select **Add**.

Example application backup schedule

Assume the following:

- Users perform database backup operations during business hours, 08:00 to 13:00.
- The automatic backups that use this policy start between 18:00 and 22:00.

In this scenario, the application backup schedule must have a start time of 0800 and a duration of 14 hours. Alternatively, the schedule can have two windows each day; one with a start time of 0800 and duration of 5 hours, and another with a start time of 1800 and a duration of 4 hours.

Table 3-5 Example settings for a NetBackup for Greenplum application backup schedule

Schedule option	Setting
Retention	2 weeks
Backup window	Sunday through Saturday 00:08:00 - 22:00:00

Configure automatic backup schedules

If you plan to have NetBackup perform automatic scheduled backups, you need to configure an automatic backup schedule.

To configure an automatic backup schedule

1. Open the policy and select the **Schedules** tab.
2. Click **Add**.
3. Specify a unique name for the schedule.
4. Select the **Type of backup**.

   See "NetBackup for Greenplum backup types" on page 16.
5. Specify the other properties for the schedule.

   See "Schedule properties" on page 15.
6. Select **Add**.

Example automatic backup schedule

Table 3-6 shows example settings for an automatic backup schedule.

Table 3-6 Example settings for a NetBackup for Greenplum automatic backup schedule

Schedule property	Setting
Retention	2 weeks
Frequency	Every week
Backup window	Sunday, 18:00:00 - 22:00:00

Add clients to a policy

The client list contains a list of the clients on which your scripts are run during an automatic backup. This list determines the clients that can send backup requests to the application schedule. A NetBackup client must be in at least one policy but can be in more than one.

For a NetBackup for Greenplum policy, clients you want to add must have the following items installed or available:

- Greenplum
- NetBackup client or server
- The backup or restore scripts

To add clients to a policy

1. Open the policy and select the **Clients** tab.
2. Select **Add**.

3 Type the name of the client and select the hardware and operating system of the client.

You need to add the coordinator, the standby coordinator, and all the segment hosts to the clients list.

4 Select **Add**.

Adding NetBackup for Greenplum scripts to the backup selections list

The backup selections list in a database policy is different from the list in non-database policies. For example, in a Standard or MS-Windows policy, the list contains files and directories to be backed up. In a database policy, you specify scripts to be run.

Add scripts to the backup selections list only if you want to create a policy for automatic backups. In that case, add the scripts to a policy that has automatic backup schedules. NetBackup runs the scripts in the order that the scripts appear in the backup selections list.

To add scripts to the backup selections list

1 Ensure that the script resides at the same location on each node. (Note that if a script is not found on a specific node, the backup runs, but the job fails for that specific node.)

See "Registering authorized locations used by a NetBackup database script-based policy" on page 37.

2 Open the policy and select the **Backup selections** tab.

3 Select **Add**.

4 In the **Pathname or directive** box, type the full path name of a script on the client.

For example:

```
/usr/openv/netbackup/ext/db_ext/greenplum_backup.sh
```

It is recommended that you copy the script to a new file name or location so that it is not overwritten during upgrades.

5 Select **Add**.

About the backup script for NetBackup for Greenplum

NetBackup for Greenplum scripts send the directives that initiate a backup of the specified database to the Greenplum server. The NetBackup for Greenplum installation process includes the following example script:

```
/usr/openv/netbackup/ext/db_ext/greenplum/scripts/greenplum_backup.sh
```

To ensure that scripts run successfully on all nodes, verify that:

- The scripts reside on each client in the client list and in the same location on each client.
- The script location is registered.
  See "Registering authorized locations used by a NetBackup database script-based policy" on page 37.
- That NetBackup can access the script location.
- The scripts reside outside of the `scripts` directory. Otherwise, subsequent NetBackup upgrades can overwrite your site's scripts.

See the following for details on how to modify the script for your environment.

See "Modify the backup script for NetBackup for Greenplum" on page 20.

Modify the backup script for NetBackup for Greenplum

The following procedure shows how to modify the backup script.

To modify the greenplum_backup.sh script

1. Copy the example script to the authorized location on your client.

   See "Registering authorized locations used by a NetBackup database script-based policy" on page 37.

   Note: Do not save custom scripts in the samples directory. Subsequent NetBackup upgrades can overwrite your site's scripts.

2. Set the access permissions of the script to 770.

   ```
   chmod 770 greenplum_backup.sh
   ```

3. Use a text editor to open the `greenplum_backup.sh` script.

   The following example uses the `vi`(1) text editor.

   ```
   vi greenplum_backup.sh
   ```

4. Locate the `USER CUSTOMIZABLE VARIABLE SECTION` and then customize the variables in the script as needed.
5. Save and close the file.
6. Test the script you modified.

   See "Perform a manual backup" on page 27.

Configuration options for backup jobs

The following table describes the values you should set for various job-related settings.

Table 3-7 Settings that control the number of jobs and how jobs run

Setting	Value
`number_of_streams`	This setting is available in the `nbgpbackup_config.yaml` file. It controls how many NetBackup jobs are created per segment host. Set the value as a multiple of the number of segments on the segment hosts. Minimum should be n * number of segments per host.
Maximum concurrent jobs per client (Primary server, Global attributes)	Set to a value that is equal to or greater than the `number_of_streams`.
Maximum concurrent jobs (storage unit)	(`number_of_streams` * *number of segment hosts*) + 1 (for the coordinator host)

Best practices for Greenplum backups

For the best backup performance all jobs need to run at the same time. For a list of configuration options, see the following topic:

See "Configuration options for backup jobs" on page 21.

You may be able to improve performance by making the following configuration changes. Make the changes in the following order:

- Start with the `gpbackup` option `--jobs 1` and the `nbgpbackup` configuration option `number_of_streams` equal to the number of segments per host.
  - The value for **Maximum jobs per client** (host property) should be set to at least the `number_of_streams`.
  - The value for **Maximum concurrent jobs** (storage unit) should be set as follows:
    (*number_of_streams* * *number of segments hosts*) + 1

- If the performance is not adequate, then increase the `buffer_size` or the `number_of_buffers`.
- If the performance still is not adequate, then you can try the following configuration:

  `--jobs = n`

  `number_of_streams` = *n * number of segments per host*

  These settings increase the number of backup jobs that `gpbackup` starts in parallel.
  - The **Maximum jobs per client** should be at least `number_of_streams`.
  - The **Maximum concurrent jobs** on the storage unit should be: (*number_of_streams * number of segment hosts*) + 1
- If the number of backup jobs is more than the value for concurrent storage unit jobs in the previous bullet, some segment hosts may perform the backup slower than others. If you increase the `end_image_timeout` value it may keep all segment files in a single image.
- If you use disk storage units, you may want to disable throttling. See the following article:
  https://support.cohesity.com/s/article/article-100005434

Use of the --jobs option or backup streams and the effect on NetBackup jobs and throttling options

The `--jobs` option for the `gpbackup` command line specifies the number of coordinator-initiated jobs that run in parallel. This setting translates to the number of worker processes that the coordinator runs concurrently. These processes send `backup_data` requests to each segment host and segment.

More information is available in the `gpbackup` utility reference guide.

The following list describes the interaction of the jobs option (`gpbackup`) and `number_of_streams` option (`nbgpbackup_config.yaml`):

- The `number_of_streams` in the `nbgpbackup_config.yaml` controls how many NetBackup jobs are created. Each NetBackup job that is created represents multiple files that are written to one image.
  The `number_of_streams` determines the number of NetBackup jobs that are created.

Controlling active NetBackup jobs

Segment hosts	Segments per host	Streams	Typical coordinator files	Total NetBackup jobs
3	4	4	5	17
10	8	8	5	85

The nbgpbackup_config.yaml file

The backup and restore commands for Greenplum require the use of a plug-in configuration file in `yaml` format. The example script creates a plug-in configuration file when the backup policy runs.

An example `yaml` file is installed with the NetBackup for Greenplum plug-in at the following location:

`/usr/openv/netbackup/ext/db_ext/greenplum/scripts/nbgpbackup_config.yaml`

If you want to run `gpbackup` or `gprestore` from the command line you must create your own `nbgpbackup_config.yaml file`.

Note: Do not save custom scripts in the `scripts` directory. Subsequent NetBackup upgrades can overwrite your site's scripts.

Supported options for the nbgpbackup_config.yaml file

The following options are supported for the `nbgpbackup_config.yaml` file if you need to customize how backups or restores run or improve performance.

Table 3-8 Options in the nbgpbackup_config.yaml file

Option	Description	Backup option	Restore option
buffer_size	The size of the shared memory buffer to use, in bytes. It should be a multiple of 512 bytes. The default is `512`. This setting also applies when you use the option `incremental_optimize_restore`.	Yes	Yes
coordinator	The host name of the coordinator host.	No	Yes

Table 3-8 Options in the nbgpbackup_config.yaml file *(continued)*

Option	Description	Backup option	Restore option
coordinator_hosts	A comma-separated list of host names. This list is used for a redirected restore to another client to find files at restore time.	No	Yes
end_image_timeout	The number of seconds for a NetBackup job to wait for another file before it ends the image. The default is `15` seconds.	Yes	No
error_shutdown_timer	The number of seconds for a `nbgpbackup` worker process to stay alive without doing anything. The default is `36000` seconds.	Yes	No
incremental_optimize_restore	Enables an optimized restore that creates a job per image. No disk cache is used. The default is `off`. When set to `on`, this option requires the `coordinator` option.	No	Yes
keyword	A keyword that can be associated with NetBackup images. Corresponds to `NBBSA_KEYWORD`.	Yes	No
number_of_buffers	The number of shared memory buffers to use per shared memory segment. The default is `30`. This setting also applies when you use the option `incremental_optimize_restore`.	Yes	Yes
number_of_streams	The number of NetBackup backup streams to use per host. See "Best practices for Greenplum backups" on page 21.	Yes	No

Table 3-8 Options in the nbgpbackup_config.yaml file *(continued)*

Option	Description	Backup option	Restore option
optimize_restore	Enables an optimized restore which caches entire images to disk so the subsequent file restores are faster. Options are `on` or `off`. The default is `off`. Requires that `optimized_restore_directory` is also specified. The NetBackup for Greenplum agent makes every attempt to clean up the `cache` directory. However, if there are errors during the restore you may have to manually clean up the cache directory on all segment hosts.	No	Yes
optimize_restore_directory	The parent directory to use for an optimized restore. The directory that is specified is created if it doesn't exist. Any parent directories are not created. For example, if `/a/b/c` is specified then `nbgpbackup` tries to create the `c` directory if it isn't present. It does not try to create `/a/b`. The directory is deleted after a successful restore.	No	Yes
policy	The NetBackup policy to use for backup. This option is required for a backup. Corresponds to `NBBSA_POLICY`.	Yes	No
read_delay	The number of milliseconds to wait before NetBackup evaluates if a shared memory buffer is full. The default is `10` milliseconds.	Yes	Yes
restore_timeout	The number of seconds for an incremental optimized restore to wait for an empty buffer. The default is `300` seconds. If you adjust the **Client read timeout** setting in the media server host properties, it is recommended that you also set the `restore_timeout` to the same value.	No	Yes
schedule	The NetBackup schedule to use for a backup. Required for a backup. Corresponds to `NBBSA_SCHEDULE`.	Yes	No

Table 3-8 Options in the nbgpbackup_config.yaml file *(continued)*

Option	Description	Backup option	Restore option
server	The NetBackup primary server to use. If not specified then the first server that is listed in the `bp.conf` is used. Corresponds to `BSA_SERVICE_HOST`.	Yes	Yes
segment_hosts	A comma-separated list of host names. This list is used for a redirected restore to another client to find files at restore time.	No	Yes
server_read_timeout	The number of seconds that `bpbrm` on the media server waits before it times out. This setting overrides the `client_read_timeout` value on the media server that is configured for `bpbrm`. The default value is `0` seconds, which instructs NetBackup to use the `client_read_timeout` value for the media server.	Yes	Yes
shutdown_process_timer	The number of seconds for any `nbgpbackup` worker processes to stay alive after they think all backup operations are complete. The default is `2` seconds.	Yes	No
write_delay	The number of milliseconds to wait before NetBackup verifies that a shared memory buffer is empty. The default is `15`.	Yes	Yes

About permissions for NetBackup for Greenplum log files

NetBackup uses the `/usr/openv/netbackup/logs` directory tree not only for the recording of troubleshooting information, but for progress and communication updates to users and other NetBackup applications. Restrictive permissions on these directories can not only disable the collection of troubleshooting data, but also prevent the application itself from functioning correctly.

See "Enable the debug logs manually (Greenplum)" on page 34.

Perform a manual backup

After you configure the servers and assets in your environment, you can test the configuration settings with a manual backup. Perform a manual backup (or backups) from a policy with the automatic backup schedules that you created.

To perform a manual backup from a policy

1. On the left, select **Protection > Policies**.
2. Select the policy you want to test.
3. Select **Manual backup**.
4. Select the schedule that you want to use for the manual backup.
5. Select the clients that you want to include for the manual backup.

Chapter 4

Using NetBackup for Greenplum

This chapter includes the following topics:

About Greenplum database backups

You can use the NetBackup interfaces to start Greenplum backups. You can also issue backup commands directly from the `gpbackup` utility to perform Greenplum backups.

The following types of backups are supported for Greenplum:

- Full backups. This type of backup copies the entire database.
- Incremental backups. This type of backup contains the database changes made since the last database backup. No schedule type is available for incremental backups. Instead, you must update the backup script to enable incremental backups. Also note that you must run a full backup before you run an incremental backup.

The most convenient way to back up your database is to set up schedules for automatic backups. When the NetBackup scheduler invokes a schedule for an automatic backup, the Greenplum backup scripts are run in the same order as they appear in the file list. The scheduler tries to find each script, and it runs the scripts that it finds. Information is available on how to initiate a manual backup or a user-directed backup of a policy.

See "Perform a manual backup" on page 27.

See “Perform a user-directed backup for Greenplum” on page 29.

Perform a user-directed backup for Greenplum

You can use the `gpbackup` command-line directly to start a Greenplum backup. NetBackup for Greenplum supports incremental backups from the incremental timestamps.

To perform a user-directed backup for Greenplum

1. Create a `yaml` file from the example file `nbgpbackup_config.yaml`.
2. On the Greenplum coordinator host run the following command:

   `gpbackup --dbname` *`database name`* `--plugin-config` *`yaml path`*

   Where *database name* is the database that you want to back up and *yaml path* is the full path of the `yaml` file.

   Refer to the documentation for `gpbackup` for additional options.

About restoring a Greenplum database

NetBackup for Greenplum supports incremental restores from the incremental timestamps.

Limitations

The following limitations apply to Greenplum database restores:

- The NetBackup for Greenplum plug-in does not support the `restore_subset` option. It should stay set to off in the `nbgpbackup_config.yaml` file.
- The plug-in does not support the `gprestore` command with the `--incremental` option.
- When you use the `optimize_restore` option, then NetBackup for Greenplum uses a cache directory. The agent makes every attempt to clean up the cache directory. However, if errors occur during a restore then you may need to manually clean up the `cache` directory on all segment hosts.

Restore options

For the best restore performance set `incremental_optimize_restore: on` in the `nbgpbackup_config.yaml` file. This setting starts a job for each image that is needed for the restore. By default, this setting is turned off and each `restore_data` request

is a separate job. For a large number of tables, in the media server host properties update the **Client read timeout** to a higher value. For example, **14400**.

Table 4-1 Effect of restore options on number of jobs

Non-optimized	optimize_restore	incremental_optimize_restore
Job for each time that `nbgpbackup` is called. This number can be thousands of jobs.	A job for each image that is used during the database restore.	A job for each image that is used during the database restore.

Table 4-2 Restore options and use of the disk cache on segment hosts

Non-optimized	optimize_restore	incremental_optimize_restore
No	Yes. Each image that is requested is restored to the `cache` directory even if a file in the image is not needed. Files are deleted after they are restored to the Greenplum database.	No

Table 4-3 Effect of restore options on speed

Non-optimized	optimize_restore	incremental_optimize_restore
Slowest	Fastest for any restores from a full backup.	Fastest for any restores from incremental backups.

Table 4-4 Restore options and use of shared memory

Non-optimized	optimize_restore	incremental_optimize_restore
No	No	Yes

Restore a Greenplum database

This topic describes how to restore to a different database.

To restore a Greenplum database

1. Create a `yaml` file from the example `nbgpbackup_config.yaml` file.
2. On the coordinator, determine which timestamp that you want to restore to. Use the `bplist` command:

   ```
   /usr/openv/netbackup/bin/bplist -l -t 24 -C client -S server
    -unix_files -R 999 /
   ```

   The timestamp is the directory name before the file. For example:

   ```
   /data/coordinator/gpseg-1/backups/20241107/20241107135750/
   gpbackup_20241107135750_report
   ```

   In the previous line, the timestamp is `20241107135750`.
3. On the coordinator, run the `gprestore` command.

   ```
   gprestore --plugin-config path_to_yaml_file --timestamp
   timestamp --redirect-db restoredb --create-db
   ```

   The value for *timestamp* should match the timestamp that you found in step 2. For example, `20241107135750`.

Redirect a restore to a different Greenplum configuration

When you redirect to a different client, you can restore the Greenplum databases to a NetBackup for Greenplum client other than the one that was originally backed up. The administrator can direct restores to any client (regardless of which client performed the backup).

Review the following requirements:

- The version of Greenplum must be the same on both clients.
- The contents of the following directory must be the same on each client: `/usr/local/greenplum-<version>/bin`
- The target Greenplum system must have the same `$GPHOME` directory structure as the source system.
- The user ID of the Greenplum database user must match that of the original host's user ID.
- The `MASTER_DATA_DIRECTORY` must be the same on both the source and the destination clients.

- The VMware Tanzu Greenplum Backup and Restore package utility and all other required utilities must be present in the following directory. This requirement applies to the coordinator and the segment hosts on the target client.

  `/usr/local/greenplum-<version>/bin`

To redirect a restore to a different Greenplum configuration

1. Ensure that the NetBackup server is configured to allow for redirected restores.

   If the `No.Restrictions` file is not used then configure a *`peername`* file in the `altnames` directory for each host that participates in the restore. Every destination segment host must have access to the images from the source segment hosts. This requirement applies to the coordinator and any mirrored coordinator and mirrored segment hosts. For more information, see the NetBackup Web UI Administrator's Guide.

2. Create a `yaml` file from the example `yaml` file.

   In this file use the `coordinator_hosts` and `segment_hosts` options to list all the source hosts that are part of the backed up Greenplum configuration.

```
coordinator_hosts: "source_cooridnator_hostname,source_mirrored_coordiantor_hostname"
segment_hosts: "source_segment_host1,source_segment_host2,source_segment_host3,…"
```

3. On the coordinator, determine which timestamp that you want to restore to. Use the `bplist` command:

   ```
   /usr/openv/netbackup/bin/bplist -l -t 24 -C client -S server
    -unix_files -R 999 /
   ```

   The timestamp is the directory name before the file. For example:

   ```
   /data/coordinator/gpseg-1/backups/20241107/20241107135750/
   gpbackup_20241107135750_report
   ```

   In the previous line, the timestamp is `20241107135750`.

4. On the coordinator, run the `gprestore` command.

   ```
   gprestore --plugin-config path_to_yaml_file --timestamp
   timestamp --redirect-db restoredb --create-db
   ```

   The value for *timestamp* should match the timestamp that you found in step 3. For example, `20241107135750`.

Chapter 5

Troubleshooting NetBackup for Greenplum

This chapter includes the following topics:

NetBackup debug logs and reports

Debug logs

The NetBackup server and client software let you enable detailed debugging logs. The information in these log files can help you troubleshoot the problems that occur outside of either the database agent or Greenplum server.

Note the following for these logs:

- These logs do not reveal the errors that occur when Greenplum server is running unless those errors also affect NetBackup. Greenplum may (or may not) write errors in the application to the NetBackup logs. Your best sources for Greenplum error information are the logs provided by Greenplum.
- Generally, each debug log corresponds to a NetBackup process and executable.

Information about the debugging log files is available in the NetBackup Troubleshooting Guide.

Also refer to the following file:

```
/usr/openv/netbackup/logs/README.debug file
```

Reports

NetBackup provides other reports that are useful in isolating problems. One such report is **All logs entries** on the server. For more information see the NetBackup Web UI Administrator's Guide.

Enable the debug logs manually (Greenplum)

This topic describes how to manually create the directories that are used for debug logging. More information is available on how on how to use logs and reports.

See the NetBackup Troubleshooting Guide.

To enable the debug logs manually

- Create the following directories on the client:

  `/usr/openv/netbackup/logs/bpbkar`

  `/usr/openv/netbackup/logs/bphdb`

  `/usr/openv/netbackup/logs/tar`
  `/usr/openv/netbackup/logs/nbgpbackup`

  For example:

  ```
  cd /usr/openv/netbackup/logs
  mkdir bphdb
  ```

About the bphdb directory on the database client

The `/usr/openv/netbackup/logs/bphdb` directory contains logs.

The following types of logs exist:

- `exts_stdout.`*`mmddyy`*

  Unless it is redirected elsewhere, NetBackup writes Greenplum script output to this file.

- `exts_stderr.`*`mmddyy`*

  Unless it is redirected elsewhere, NetBackup writes Greenplum script errors to this file.

- `log.`*`mmddyy`*

  This log contains debugging information for the `bphdb` process. `bphdb` is the NetBackup database backup binary. It is invoked when an automatic backup

schedule is run. NetBackup for Greenplum uses this client process for Greenplum script execution.

Set the debug level on a Greenplum client

To control the amount of information that is written to the debug logs, change the "Database" debug level. Typically, the default value of 0 is sufficient. However, Technical Support may ask you to set the value higher to analyze a problem.

The debug logs are located in `/usr/openv/netbackup/logs`.

To set the debug level

- Enter the following line in the `bp.conf` file.

  `VERBOSE = X`

  Where *X* is the debug level you want.

Minimize time-out failures on large database restores

Large database restores sometimes fail when multiple restore sessions compete for resources. In this situation, a restore session can be delayed while NetBackup waits for media or device access. If the delay is too long, the restore session times out. Use the following procedure to minimize session time-outs and to allow the restore jobs to complete successfully.

To minimize time out failures on large database restores

1. Open the NetBackup web UI.
2. On the left, select **Hosts > Host properties**.
3. Select the check box for the client.
4. If necessary, select **Connect**.
5. Select **Edit client**.
6. Select **Timeouts**.

7 Set the **Client read timeout** property to a large value.

The default for the **Client read timeout** setting is 300 seconds (5 minutes). For database agent clients, increase the value significantly from the recommended value.

See the NetBackup Web UI Administrator's Guide.

For example, change this setting to 30-60 minutes to minimize time-out errors.

8 Select **Save**.

Note: This change may delay detecting problems during subsequent backups. Consider putting the original value back in place once any restore that requires a change is complete.

Troubleshooting of Greenplum backups that are canceled mid-backup

If Greenplum backups are canceled mid-backup, the backup processes are left hanging. If the main Greenplum `gpbackup` process is terminated in mid-backup, the `cleanup_plugin_for_backup` calls cannot be made to the plug-in to complete the normal completion and shutdown of the backup. In this case the `setup_plugin_for_backup` processes hang that remain running on the segment hosts. These processes eventually end due to the configurable `error_shutdown_timer` setting in the `nbgpbackup_config.yaml` file. The default is 10 hours.

Appendix A

Register authorized locations

This appendix includes the following topics:

Registering authorized locations used by a NetBackup database script-based policy

During a backup, NetBackup checks for scripts in the default script location and any authorized locations. The default, authorized script location for UNIX is `usr/openv/netbackup/ext/db_ext` and for Windows is *`install_path`*`\netbackup\dbext`. If the script is not in the default script location or an authorized location, the policy job fails. You can move any script into the default script location or any additional authorized location and NetBackup recognizes the scripts. You need to update the policy with the script location if it has changed. An authorized location can be a directory and NetBackup recognizes any script within that directory. An authorized location can also be a full path to a script if an entire directory does need to be authorized.

If the default script location does not work for your environment, use the following procedure to enter one or more authorized locations for your scripts. Use `nbsetconfig` to enter an authorized location where the scripts reside. You can also use `bpsetconfig`, however this command is only available on the primary or the media server.

Note: One recommendation is that scripts should not be world-writable. NetBackup does not allow scripts to run from network or remote locations. All scripts must be stored and run locally. Any script that is created and saved in the NetBackup `db_ext` (UNIX) or `dbext` (Windows) location needs to be protected during a NetBackup uninstall.

For more information about registering authorized locations and scripts, review the knowledge base article:

https://support.cohesity.com/s/article/article-100039639

To add an authorized location

1. Open a command prompt on the client.
2. Use `nbsetconfig` to enter values for an authorized location. The client privileged user must run these commands.

   The following examples are for paths you may configure for the Oracle agent. Use the path that is appropriate for your agent.

   - On UNIX:

     ```
     [root@client26 bin]# ./nbsetconfig
     nbsetconfig>DB_SCRIPT_PATH = /Oracle/scripts
     nbsetconfig>DB_SCRIPT_PATH = /db/Oracle/scripts/full_backup.sh
     nbsetconfig>
     <ctrl-D>
     ```

   - On Windows:

     ```
     C:\Program Files\Cohesity NetBackup\NetBackup\bin>nbsetconfig
     nbsetconfig> DB_SCRIPT_PATH=c:\db_scripts
     nbsetconfig> DB_SCRIPT_PATH=e:\oracle\fullbackup\full_rman.sh
     nbsetconfig>
     <ctrl-Z>
     ```

Note: Review the NetBackup Command Reference Guide for options, such as reading from a text file and remotely setting clients from a NetBackup server using `bpsetconfig`. If you have a text file with the script location or authorized locations listed, `nbsetconfig` or `bpsetconfig` can read from that text file. An entry of `DB_SCRIPT_PATH=none` does not allow any script to run on a client. The `none` entry is useful if an administrator wants to completely lock down a server from running scripts.

3 (Conditional) Perform these steps on any clustered database or agent node that can perform the backup.

4 (Conditional) Update any policy if the script location was changed to the default or authorized location.