NetBackup™ Backup Planning and Performance Tuning Guide

Release 8.3 and later

Document version 9

VERITAS™

NetBackup™ Backup Planning and Performance Tuning Guide

Last updated: 2024-04-16

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

NetBackup capacity planning

This chapter includes the following topics:

Purpose of this guide

This guide covers NetBackup release 8.3 and later. The purpose of this guide is to provide recommendations and guardrails based on extensive field experience. Results may differ based on your unique environment.

Veritas NetBackup is a high-performance data protection application. Its architecture is designed for large and complex distributed computing environments. NetBackup provides scalable storage servers (primary and media servers) that can be configured for network backup, recovery, archiving, and file migration services.

This guide is for administrators who want to analyze, evaluate, and tune NetBackup performance. It is intended to provide guidance on questions such as the following: How big should the NetBackup primary server be? How can the server be tuned for maximum performance? How many CPUs and disk or tape drives are needed?

How to configure backups to run as fast as possible? How to improve recovery times? What tools can characterize or measure how NetBackup handles data? How to tune NetBackup for different workloads, in particular with a high number of small concurrent jobs and a few large jobs.

Note: Most critical factors in performance are based in hardware rather than software. Compared to software, hardware configuration has roughly four times greater effect on performance. Although this guide provides some hardware configuration assistance, it is assumed for the most part that your devices are correctly configured.

For additional planning or performance-related information, refer to the following documents:

- *Support for NetBackup 7.7.x, 8.x, 9.x, and 10.x in virtual environments*
  https://www.veritas.com/content/support/en_US/doc/NB_70_80_VE
- *NetBackup Deduplication Guide*
  Go to the NetBackup Documentation Landing Page, select the appropriate release, and then the guide:
  https://www.veritas.com/content/support/en_US/article.100040135
- *Storage Lifecycle Policy (SLP) Cheat Sheet*
  https://www.veritas.com/content/support/en_US/article.100006475

Changes in Veritas terminology

To modernize our terminology, Veritas has begun to replace certain outdated terms with more current terms.

Note: As Veritas continues to update its terminology, the deprecated terms and the new terms may be used interchangeably.

Deprecated term	New term
Master	Primary
Slave	Secondary or media server
Whitelist or white list	Allowed list
Blacklist or black list	Blocked list
White hat	Ethical

Deprecated term	New term
Black hat	Unethical

Disclaimer

It is assumed you are familiar with NetBackup and your applications, operating systems, and hardware. The information in this manual is advisory only, presented in the form of guidelines. Changes to an installation that are undertaken as a result of the information in this guide should be verified in advance for appropriateness and accuracy. Some of the information that is contained herein may apply only to certain hardware or operating system architectures.

Note: The information in this manual is subject to change.

How to analyze your backup requirements

Many factors can influence your backup strategy. You should analyze these factors and then make backup decisions according to your site's priorities.

When you plan your installation's NetBackup capacity, ask yourself the following questions:

Table 1-1 Questions to ask as you plan NetBackup capacity

Questions	Actions and related considerations
Which systems need to be backed up?	Identify all systems that need to be backed up. List each system separately so that you can identify any that require more resources to back up. Document which computers have disk drives or tape drives or libraries attached and write down the model type of each drive or library. Identify any applications on these computers that need to be backed up, such as Oracle, DB2, VMware, MySQL, or MS-Exchange. In addition, record each host name, operating system and version, database type and version, network technology (for example, 10 gigabits), and location.

Table 1-1 Questions to ask as you plan NetBackup capacity *(continued)*

Questions	Actions and related considerations
How much data is to be backed up?	Calculate how much data you need to back up and the daily/weekly/monthly/yearly change rate. The change rates affect the deduplication ratio and therefore the amount of data that need to be written to the disk group. Include the total disk space on each individual system, including that for databases. Remember to add the space on mirrored disks only once. By calculating the total size of all clients, you can design a system that takes future growth into account. Try to estimate how much data you will need to back up in 6 months to a few years from now. Consider the following: ■ Do you plan to back up databases or raw partitions? To back up databases, identify the database engines, their version numbers, and the method to back them up. NetBackup can back up several database engines and raw file systems, and databases can be backed up while they are online or offline. To back up a database that is online, you need a NetBackup database agent for your particular database engine. With a Snapshot Client backup of databases using raw partitions, you back up as much data as the total size of your raw partition. Also, remember to add the size of your database backups to your final calculations when figuring out how much data you need to back up. ■ Do you plan to back up special application servers such as MS-Exchange or Lotus Notes? To back up application servers, identify their types and application release numbers. As previously mentioned, you may need a special NetBackup agent to back up your particular servers.
Should Accelerator be enabled for VMware virtual machine backup?	In subsequent full backups and incremental backups, only changed data would be backed up when Accelerator is enabled in the NetBackup policy. This case would greatly shorten the backup time and reduce the backup size. It's strongly recommended enabling Accelerator for VMware VMs backup. ■ With Accelerator for VMware, the type of data matters less. What matters the most is the change rate, (how much data changes from one full to the next). In most typical workloads, the change rate is quite small, therefore a large benefit can be achieved using Accelerator. ■ The biggest benefit of enabling Accelerator is realized for full backup schedules. A full backup image is created at the cost of an incremental (that is, time and data transfer). ■ An incremental schedule backup for VMware with Accelerator produces a full image of the VM, but only changed files are cataloged. This approach helps in use cases such as Instant Restore/Instant Access can be done from an incremental image as well as simplifies optimized duplication (opt-dup) replication. Tape-out and non-opt-dup replication can have an adverse effect, as the size of such an image can be as big as a full image.

Table 1-1 Questions to ask as you plan NetBackup capacity *(continued)*

Questions	Actions and related considerations
What types of backups are needed and how often should they take place?	The frequency of your backups has a direct effect on your: ■ Disk or tape requirements ■ Data transfer rate considerations ■ Restore opportunities. To properly size your backup system, you must decide on the type and frequency of your backups. Will you perform daily incremental and weekly full backups? Monthly or bi-weekly full backups?
How much time is available to run each backup?	What is the window of time that is available to complete each backup? The length of a window dictates several aspects of your backup strategy. For example, you may want a larger window of time to back up multiple, high-capacity servers. Or you may consider the use of advanced NetBackup features such as synthetic backups, a local snapshot method, or FlashBackup.
Should the scheduling windows for backups overlap with those of duplication or replication jobs or should they be separated?	If the windows of backup and duplication or replication jobs (including Auto Image Replication (A.I.R.)) overlap, performance of these jobs would be affected. Carefully design the scheduling of backups and duplication or replication jobs to try to avoid overlapping. For more information, see the following documents: *Auto Image Replication (A.I.R.) slow performance, particularly for small images*: https://www.veritas.com/content/support/en_US/article.100045506 *How to tune NetBackup Auto Image Replication (A.I.R.) operations for maximum performance*: https://www.veritas.com/content/support/en_US/article.100046559
Is archiving to the cloud supported?	NetBackup supports various cloud archive technologies including AWS Glacier options, Snowball, and Snowball Edge, along with Microsoft Azure Archive.
How long should backups be retained?	An important factor while you design your backup strategy is to consider your policy for backup expiration. The amount of time a backup is kept is also known as the retention period. A fairly common policy is to expire your incremental backups after one month and your full backups after 6 months. With this policy, you can restore any daily file change from the previous month and restore data from full backups for the previous 6 months. The length of the retention period depends on your own unique requirements and business needs, and perhaps regulatory requirements. However, the length of your retention period is directly proportional to the number of tapes you need and the size of your NetBackup image database. Your NetBackup image database keeps track of all the information on all your disk drives and tapes. The image database size is tightly tied in to your retention period and the frequency of your backups.
If backups are sent off site, how long must they remain off site?	If you plan to send tapes off-site as a disaster recovery option, identify which tapes to send off site and how long they remain off-site. You might decide to duplicate all your full backups, or only a select few. You might also decide to duplicate certain systems and exclude others. As tapes are sent off site, you must buy new tapes to replace them until they are recycled back from off-site storage. If you forget this detail, you may run out of tapes when you most need them.

Table 1-1 Questions to ask as you plan NetBackup capacity *(continued)*

Questions	Actions and related considerations
What is your network technology?	If you plan to back up any system over a network, note the network types. The next section explains how to calculate the amount of data you can transfer over those networks in a given time. Based on the amount of data that you want to back up and the frequency of those backups, consider using 10-GB network interfaces, linking aggregation/teaming, or installing a private network for backups.
What systems do you expect to add in the next 6 months?	Plan for future growth when you design your backup system. Analyze the potential growth of your system to ensure that your current backup solution can accommodate future requirements. Remember to add any resulting growth factor that you incur to your total backup solution.
Will user-directed backups or restores be allowed?	Allow users to do their own backups and restores, to reduce the time it takes to initiate certain operations. However, user-directed operations can also result in higher support costs and the loss of some flexibility. User-directed operations can monopolize disk pools and tape drives when you most need them. They can also generate more support calls and training issues while the users become familiar with the new backup system. Decide whether user access to some of your backup systems' functions is worth the potential cost.
What data types are involved?	What are the types of data: text, graphics, database, virtual machines? How compressible is the data? What is the typical deduplication rate of data to be backed up? How many files are involved? Will NetBackup's Accelerator feature be enabled for VMware virtual machine or NDMP backups? (Note that only changed data is backed up with Accelerator for both full and incremental backup.) Will the data be encrypted? (Note that encrypted backups may run slower.)
Where is the data located?	Is the data local or remote? Is it tape, JBOD (just a bunch of disks), or disk array? What are the characteristics of the storage subsystem? What is the exact data path? How busy is the storage subsystem?
How to test the backup system?	Because hardware and software infrastructure can change over time, create an independent test backup environment. This approach ensures that your production environment can work with the changed components.

How to calculate the size of your NetBackup image database

An important factor when you design your backup system is to calculate how much disk space is needed to store your NetBackup image database. Your image database keeps track of all the files that have been backed up.

The image database size depends on the following variables, for both full backups and incremental backups:

- The number of files being backed up
- The frequency and the retention period of the backups

You can use either of two methods to calculate the size of the NetBackup image database. In both cases, since data volumes grow over time, you should factor in expected growth when calculating total disk space used.

NetBackup can be configured to automatically compress the image database to reduce the amount of disk space required. When a restore is requested, NetBackup automatically decompresses the image database, only for the time period needed to accomplish the restore. You can also use archiving to reduce the space requirements for the image database. More information is available on catalog compression and archiving.

See the *NetBackup Administrator's Guide, Volume I*.

Note: If you select NetBackup's true image restore option, your image database becomes larger than an image database without this option selected. True image restore collects the information that is required to restore directories to their contents at the time of any selected full or incremental backup. The additional information that NetBackup collects for incremental backups is the same as the information that is collected for full backups. As a result, incremental backups take much more disk space when you collect true image restore information.

First method: You can use this method to calculate image database size precisely. It requires certain details: the number of files that are held online and the number of backups (full and incremental) that are retained at any time.

To calculate the size in gigabytes for a particular backup, use the following formula:

Image database size = (132 * number of files in all backups)/ 1GB

To use this method, you must determine the approximate number of copies of each file that is held in backups and the typical file size. The number of copies can usually be estimated as follows:

Number of copies of each file that is held in backups = number of full backups + 10% of the number of incremental backups held

The following is an example of how to calculate the size of your NetBackup image database with the first method.

This example makes the following assumptions:

- Number of full backups per month: 4
- Retention period for full backups: 6 months
- Total number of full backups retained: 24

- Number of incremental backups per month: 25
- Retention period for incremental backups per month: 1 month
- Total number of files that are held online (total number of files in a full backup): 17,500,000

Solution:

Number of copies of each file retained:

24 + (25 * 10%) = 26.5

NetBackup image database size for each file retained:

(132 * 26.5 copies) = 3498 bytes

Total image database space required:

(3498 * 17,500,000 files) /1 GB = 61.2 GB

Second method: Multiply by a small percentage (such as 2%) the total amount of data in the production environment (not the total size of all backups). Note that 2% is an example; this section helps you calculate a percentage that is appropriate for your environment.

Note: You can calculate image database size by means of a small percentage only for environments in which it is easy to determine the following: the typical file size, typical retention policies, and typical incremental change rates. In some cases, the image database size that is obtained using this method may vary significantly from the eventual size.

To use this method, you must determine the approximate number of copies of each file that are held in backups and the typical file size. The number of copies can usually be estimated as follows:

Number of copies of each file that is held in backups = number of full backups + 10% of the number of incremental backups held

The multiplying percentage can be calculated as follows:

Multiplying percentage = (132 * number of files that are held in backups / average file size) * 100%

Then, the size of the image database can be estimated as:

Size of the image database = total disk space used * multiplying percentage

The following is an example of how to calculate the size of your NetBackup image database with the second method.

This example makes the following assumptions:

- Number of full backups per month: 4
- Retention period for full backups: 6 months
- Total number of full backups retained: 24
- Number of incremental backups per month: 25
- Retention period for incremental backups per month: 1 month
- Average file size: 70 KB
- Total disk space that is used on all servers in the domain: 1.4 TB

Solution:

Number of copies of each file retained:

24 + (25 * 10%) = 26.5

NetBackup image database size for each file retained:

(132 * 26.5 copies) = 3498 bytes

Multiplying percentage:

(3498/70000) * 100% = 5%

Total image database space required:

(1,400 GB * 5%) = 70 GB

Sizing for capacity with MSDP

This section provides guidelines for sizing storage when using NetBackup Media Server Deduplication Pools (MSDP). Capacity sizing is always dependent on the nature of the data being protected in the environment and can vary significantly between environments. This section is for basic sizing methodology guidance. Assistance with sizing is available via your Veritas account team or reseller.

Key sizing parameters

Sizing depends on answering several key questions:

- What kind of data is being protected and how well does it deduplicate?
- How much data is being protected?
- How long is it being kept and how is it being backed up?
- What is the change rate?

Data types and deduplication

Different data types deduplicate at different rates. MSDP performs both deduplication and compression of data. Deduplication is performed first and then the resulting data segments are compressed before they are written to disk.

Unstructured data

It is important to understand the different types of unstructured data in the environment for sizing. Some data types will not deduplicate well:

- Encrypted files:
  Encrypted files will not deduplicate well, and even small changes will often change the entire file resulting in higher change rates than non-encrypted files. There will generally only be small (<10% at best) storage savings from compression. There will be no deduplication between files, which will lower deduplication rates.
- Compressed, image, audio, and video files:
  Files that fall into this category will not deduplicate well, and there will be no savings from compression.

Note that encryption and compression at the file system level such as with NTFS is transparent to NetBackup, as the files are uncompressed and decrypted by the operating system when they are read. This may result in backups appearing larger in FETB than the data consumed on the file system. These file systems will see good deduplication and compression rates when the data is written to MSDP however.

Databases

Database deduplication will generally be lower than that observed for unstructured data. To achieve optimal deduplication, compression and encryption should not be enabled in the backup stream (for example, with RMAN directives for Oracle).

Database transaction logs will not deduplicate well due to the nature of the data, although savings from compression may be observed. it is important to determine deduplication rates for database backups and transaction log backups separately.

Transparent database encryption options will lower deduplication and compression rates. Initial backups will show minimal space savings. The level of deduplication achieved between backups depends on the nature of the changes to the database. In general, OLTP databases that may have changes distributed throughout the database will show lower deduplication rates than OLAP instances which tend to have more inserts than updates.

NDMP

The notes above for unstructured data apply to NDMP backups. In addition, the nature of NDMP can affect deduplication rates. NDMP defines the communication protocol between filers and backup targets. It does not define the data format. Veritas has developed stream handlers for several filers (NetApp and Dell EMC PowerScale) that allow an understanding of the data streams. Filers without a stream handler may show very low deduplication rates (for example, 20% or lower). In these cases, MSDP Variable Length Deduplication (VLD) should be enabled on the MSDP policies, and a significant increase in deduplication rates will generally be observed.

Virtualization

For virtualization workloads, supported file systems and volume managers should be used so that NetBackup can understand the structure of the data. On configurations that meet these requirements, the deduplication engine will respect file boundaries when segmenting the data stream and significant increases in deduplication rates will be observed.

Determining deduplication rates

Due to the wide variations in customer environments, even within specific workloads, Veritas does not publish expected deduplication rates.

It is recommended that customers perform test in their own environments with a representative subset of data to be protected to determine the actual deduplication rates for the schedule types to be implemented:

- Initial Full
- Daily Differential
- Subsequent Full
- Database Transaction Log

Deduplication rates can be found in the Activity Monitor in the **Deduplication Rate** column. When viewing the job details, there is also an entry for deduplication rates:

```
Oct 8, 2021 12:22:20 AM - Info media-server.example.com (pid=29340)
 StorageServer=PureDisk:mediaserver.example.com; Report=PDDO Stats
 (multi-threaded stream used) for (mediaserver.example.com): scanned:
 1447258 KB, CR sent: 6682 KB, CR sent over FC: 0 KB, dedup: 99.5%,
 cache hits: 11263 (99.2%), where dedup space saving:99.2%, compression
 space saving:0.3%
```

In this example, the deduplication rate that will be used for calculations is the total rate of 99.5%, which includes savings from compression.

Tests should be run over a period of weeks to capture typical change rates in the environment.

Determining FETB for workloads

The Front-End Terabytes (FETB) is the amount of data protected by NetBackup. Note that this is not the space allocated to clients, but the size of the data that NetBackup will protect. For example, a VMware virtual machine may have an 80 GB VMDK, but only 25GB is used, and after swap and paging files are excluded NetBackup reports 21 GB as the size of the backup. For sizing, 21 GB should be used as the FETB.

Database transaction logs need to be handled differently than other workloads, since the change rate for each backup will be 100%. For transaction logs, determine the total volume of logs that will be protected during the retention period.

Retention periods

To size any backup storage, how long data is kept must be known. Some thought should be put into defining retention periods. Important considerations include:

- What regulatory requirements exist for data retention that must be met by the backup system.
- What SLAs are present.
- How long does data need to be retained to properly recover from a ransomware attack. Ransomware often installs in the environment and sits dormant for a period, which should be considered.

Change rate

Data change rate will determine how much new data is written each backup cycle. The best way to determine this is via the NetBackup Activity Monitor and examining the size of differential backups.

When sizing for long-term retention, that nature of the change is important to understand. Consider the case of a 5% daily change rate. When the month one backup is compared to the month two backup, the change between these backups can range from 5% (the case where the same data changes every day) to 100% (the case where different data changes every day). In most environments the former is more likely than the latter. There are cases for the latter, such as security video storage.

Replication and duplication of backups

When replicating or duplicating images between pools, data commonality should be considered. Consider the case of two sites replicating bi-directionally. Each site has clients that are protected and sends those backups to the other site. If there is common data between the sites (for example, copies of databases), the data will not have to be written at the remote site, since it already exists in the storage pool. This can result in significant space savings depending on the level of data commonality. If there are large quantities of common data, this should be considered when sizing. If there is limited or no commonality, then the space consumed on the remote pool will be about the same as the space consumed locally for those backups.

Sizing calculations for MSDP clients

For a given client, the amount of storage consumed is:

$FETB*(1-D_I) + FETB*(R_F-1)*(1-D_S) + (FETB*C_D)*(1-D_D)*(R_D)$

Where:

- FETB is the front-end terabytes.
- D_I is the initial full deduplication rate.
- D_S is the subsequent full deduplication rate.
- D_D is the daily deduplication rate.
- C_D is the daily change rate.
- R_F is the total number of full backups to be retained.
- R_D is the total number of incremental/differential backups to be retained.

Note: More information about sizing for MSDP is available:

See “MSDP sizing considerations” on page 65.

About how to design your OpsCenter server

NetBackup OpsCenter is a web-based software application that provides detailed information on your data protection environment. It can track the effectiveness of data backup and archive operations by generating comprehensive reports.

For assistance in planning and designing an OpsCenter installation, refer to the following documents:

- *NetBackup OpsCenter Administrator's Guide*

- *NetBackup OpsCenter Performance and Tuning Guide*

Chapter 2

Primary server configuration guidelines

This chapter includes the following topics:

Size guidance for the NetBackup primary server and domain

NetBackup primary server sizing is an important activity as part of an overall NetBackup solution design. Veritas always recommends a comprehensive data protection assessment to determine the optimal configuration for a NetBackup primary and NetBackup domain.

The following information is meant as guidelines:

- NetBackup has no hard limit on catalog size. However, Veritas recommends as a best practice that you keep the catalog size under 4 TB to ensure good catalog backup and recovery performance.
  The size of the NetBackup catalog and the performance that is related to reading data from the NetBackup catalog is driven by the I/O performance and more specifically the disk speed. Veritas recommends the use of solid-state drives (SSDs) where possible for the catalog. The disks require good read and write performance, which is even more critical in large environments.
  Managing the size of the catalog through compression and catalog archiving is recommended for images with a long-term retention (LTR).
  See “Methods for managing the catalog size” on page 38.
- The number of devices in the EMM database should not exceed 1500.
  Examples of devices are a tape drive, a tape library, a disk pool, and so on.
- The number of media servers should not exceed 50.
  It is important to maintain a manageable number of media servers and storage targets within each NetBackup domain. Every media server and storage target that is deployed must be managed, maintained, and eventually patched and upgraded. Each of those media servers has a configuration that has to also be maintained. Therefore, it is important to consider the manageability, usability, and the administrative implications. Veritas recommends deploying media servers and storage targets that are properly sized with the necessary CPU, memory, network bandwidth, and disk I/O to support the backup workloads. It is also important to consider whether the same workloads require duplication or replication to a DR location. Sizing the media servers and storage targets to accommodate those secondary options is crucial. In summary, Veritas recommends that you deploy properly sized media servers and storage targets, while keeping the number less than 50 per domain.
- The number of jobs must not exceed one job per second per client, but it is possible to submit multiple jobs per second, each sent from a different client. Each backup client has the "one job per second per client" limit, so multiple clients may run in parallel.

- Computing resources such as CPU and memory affect how well the primary server scales.

To accommodate the processing of the metadata streams from media servers, it is critical that the primary server has the requisite amount of system resources. A media server sends metadata about the files it has backed up to the primary server. This metadata is batched and sent periodically. The batch size, which is determined by the tuning parameter `MAX_ENTRIES_PER_ADD`, has significant effect on primary server performance, especially for backup images that contain many small files.

See "Adjusting the batch size for sending metadata to the NetBackup catalog" on page 37.

The primary server must then process each of these metadata message payloads. Each payload requires an operating system process, each of which consumes system resources. The consumed system resources are disk capacity, CPU cycles, memory capacity, network bandwidth, and disk I/O.

Table 2-1 provides additional information.

Table 2-1 Sizing guidelines

Number of processors	Recommended memory requirement	Maximum number of media servers per primary server *
8	128 GB	20
16	256 GB	100

*Veritas recommends that you limit the number of media servers to less than 50 media servers per domain.

Factors that limit job scheduling

When many requests are submitted to NetBackup simultaneously, NetBackup increases its use of memory. The number of requests may eventually affect the overall performance of the system. This type of performance degradation is associated with the way a given operating system handles memory requests. It may affect the functioning of all applications that currently run on the system, not limited to NetBackup.

Note: In the NetBackup Administration Console, the Activity Monitor may not update if there are thousands of jobs to view. In this case, you may need to change the memory setting by means of the NetBackup Java command `jnbSA` with the `-mx` option. See the "INITIAL_MEMORY, MAX_MEMORY" subsection in the *NetBackup Administrator's Guide, Volume I*. Note that this situation does not affect NetBackup's ability to continue running jobs.

See "NetBackup job delays" on page 28.

More than one backup job per second

Starting from NetBackup 8.2, NetBackup removed the one backup job per second limitation within a primary server. With the change, the limit of starting only one job per second still holds for a single client, however, multiple jobs from different clients may be started within a second. NetBackup can scale to higher job counts with appropriate hardware resources.

Multiple backup jobs from multiple clients all starting at the same time may cause a temporary CPU and memory usage surge, sometimes significantly. The overall performance effect maybe marginal, however, if primary server is already CPU-bound or memory-bound, this temporary surge can cause the system to become unresponsive.

Stagger the submission of jobs for better load distribution

When the backup window opens, Veritas recommends scheduling jobs to start in small groups periodically, rather than starting all jobs at the same time. If the submission of jobs is staggered, the NetBackup processes can more rapidly process and allocate resources to the jobs. In addition, job staggering can help prevent server resource exhaustion.

The best job scheduling depends on many factors, such as workload type, backup type, and balancing those factors against available resources. It is recommended that performance metrics be reviewed periodically to make the necessary adjustments to maintain optimal job management.

See "NetBackup job delays" on page 28.

NetBackup job delays

NetBackup jobs may be delayed for a variety of reasons. Common delays that may occur, and in some cases suggests possible remedies, are described below..

Delays in starting jobs

The NetBackup Policy Execution Manager (`nbpem`) may not begin a backup at exactly the time a backup policy's schedule window opens. This delay can happen when you define a schedule or modify an existing schedule with a window start time close to the current time.

For instance, suppose that you create a schedule at 5:50 P.M., and specify that at 6:00 P.M. backups should start. You complete the policy definition at 5:55 P.M. At 6:00 P.M., you expect to see a backup job for the policy start, but it does not. Instead, the job takes another several minutes to start.

The explanation is the following: NetBackup receives and queues policy change events as they happen, but processes them periodically as configured in the **Policy Update Interval** setting. (The **Policy Update Interval** is set under **Host Properties > Primary Server > Properties > Global Settings**. The default is 10 minutes.) The backup does not start until the first time NetBackup processes policy changes after the policy definition is completed at 5:55 P.M. NetBackup may not process the changes until 6:05 P.M. For each policy change, NetBackup determines what needs to be done and updates its work list accordingly.

Delays in running queued jobs

Note: For any one client, there is a limit of starting only one job per second . However, multiple jobs from different clients can be started within the same second. Depending on the configuration involved, significantly more backup jobs can be started within any fixed time period, which may change the performance behavior of the system. In some environments, you may need to change some configuration settings to achieve the optimum performance behavior.

If jobs are queued and only one job runs at a time, use the **State Details** column in the Activity Monitor to see the reason for the job being queued.

If jobs are queued and only one job runs at a time, set one or more of the following to allow jobs to run simultaneously:

- **Host Properties > Primary Server > Properties > Global Attributes > Maximum jobs per client** (should be greater than 1).
- **Media and Device Management > Disk Pools > Limit I/O streams per volume**

Select to limit the number of read and write streams (that is, jobs) for each volume in the disk pool. A job may read or write backup images. If you select this property, you also need to configure the number of streams to allow per volume. When the limit is reached, NetBackup chooses another volume for write operations, if available. If not available, NetBackup queues jobs until a volume is available. In a product like Netbackup Flexible Scale (NBFS), this property is unselected by default. (that is, no limit).

- Policy attribute **Limit jobs per policy** (should be greater than 1).
- Schedule attribute **Media multiplexing** (should be greater than 1).
  The general recommendation is to start with a value of 4 and gradually increase it until you find an acceptable balance for both backup and restores. See *Media multiplexing (schedule attribute)* In the *NetBackup Administrator's Guide, Volume 1*, for more information on Media multiplexing

  Note: Keep in mind that increasing this value may affect restore times. More information is available:

  See "How fragment size affects restore of a multiplexed image on tape" on page 206.

  See "Other NetBackup restore performance issues" on page 208.

- Check the following storage unit properties:
  - Is the storage unit enabled to use multiple tape drives (**Maximum concurrent write drives**)? If you want to increase this value, remember to set it to fewer than the number of tape drives available to this storage unit. Otherwise, restores and other non-backup activities cannot run while backups to the storage unit are running.
  - Is the storage unit enabled for multiplexing (**Maximum streams per drive**)? You can write a maximum of 32 jobs to one tape at the same time.

    Note: Values greater than 4 may actually decrease overall performance because they may reduce restore speeds. More information is available:

    See "How fragment size affects restore of a multiplexed image on tape" on page 206.

    See "Other NetBackup restore performance issues" on page 208.

Note: Be sure to check with the disk storage manufacturers for recommendations about your disk storage units.

For example, you can run multiple jobs to a single storage unit if you have multiple drives. (Maximum concurrent write drives set to greater than 1.) Or, you can set up multiplexing to a single drive if Maximum streams per drive is set to greater than 1. If both Maximum concurrent write drives and Maximum streams per drive are greater than 1: you can run multiple streams to multiple drives, assuming that Maximum jobs per client is set high enough.

Note: All storage units that reference a specific MSDP pool have a maximum concurrent jobs setting. The total number of concurrent jobs for all storage units accessing a single MSDP pool should be approximately 10% less than the maximum number of I/O streams on the disk pool. This allows for secondary operations, like replications and duplications, as well as restores to be executed even during busy backup windows.

Delays in tape jobs becoming active

Tape jobs become active as soon as the resources are allocated.

NetBackup makes the tape jobs active as follows:

- The NetBackup Job Manager (`nbjm`) requests resources from the NetBackup Resource Broker (`nbrb`) for the job.
- `nbrb` allocates the resources and gives the resources to `nbjm`.
- `nbjm` starts `bpbrm` which in turn starts `bptm`.

Job delays caused by unavailable media

The job fails if no other storage units are usable, in any of the following circumstances:

- If the media in a storage unit are not configured or are unusable (such as expired)
- The maximum mounts setting was exceeded
- The wrong pool was selected

If media are unavailable, consider the following:

- Add new media
- Or change the media configuration to make media available (such as changing the volume pool or the maximum mounts).

If the media in a storage unit are usable but are busy, the job is queued. In the NetBackup Activity Monitor, the "State Details" column indicates why the job is queued, such as "media are in use." (The same information is available in the Job Details display. Right-click on the job and select "Details.") If the media are in use, the media eventually stop being used and the job runs.

Job delays after removing a media server

A job may be queued by the NetBackup Job Manager (`nbjm`) if the media server is not available. The job is not queued because of communication timeouts, but because EMM knows that the media server is down and the NetBackup Resource Broker (`nbrb`) queues the request to be retried later.

If no other media servers are available, EMM queues the job under the following circumstances:

- If a media server is configured in EMM but the server has been physically removed, turned off, or disconnected from the network
- If the network is down

The Activity Monitor should display the reason for the job queuing, such as "media server is offline." Once the media server is online again in EMM, the job starts. In the meantime, if other media servers are available, the job runs on another media server.

If a media server is not configured in EMM, regardless of the physical state of the media server, EMM does not select that media server for use. If no other media servers are available, the job fails.

To permanently remove a media server from the system, consult the "Decommissioning a media server" section in the *NetBackup Administrator's Guide, Volume I*.

Selection of storage units: performance considerations

Many different NetBackup mechanisms write backup images to storage devices, such as: backup policies, storage lifecycle policies (SLPs), staging storage units, Vault duplication, and ad hoc (manual) duplication. When writing a backup image to storage, you can tell NetBackup how to select a storage unit or let NetBackup choose the storage unit.

There are three ways to specify destination for backups and duplications:

- Storage unit (the preferred way)
- Storage Unit Group (slower than storage units)

If tape storage unit groups are used, limit the number of storage units in a storage unit group to 5 or less.

Note: Use of storage unit groups is not recommended with MSDP pools. It is recommended to specify a specific set of clients or specific workload to a consistent and specific MSDP pool. This optimizes capacity and deduplication efficacy.

- Any Available.
  Use only in test environments.

Although Storage Unit is the preferred method for most large environments, the following sections, Performance considerations for the Any Available method and Performance considerations for the Storage Unit Groups method discuss the pros and cons of specifying a storage unit group versus allowing NetBackup to choose from a group (Any Available).

Note: The more narrowly defined the storage unit designation is, the faster NetBackup can assign a storage unit, and the sooner the job starts.

Performance considerations for the Any Available method

As a rule, the Any Available method should only be used in small, simple environments.

For most backup operations, the default is to let NetBackup choose the storage unit (a storage destination of Any Available). Any Available may work well in small configurations that include relatively few storage units and media servers.

However, Any Available is NOT recommended for the following:

- Configurations with many storage units and media servers. Any Available is not recommended.
- Configurations with disk technologies (such as AdvancedDisk, PureDisk, OpenStorage). With these newer disk technologies, Any Available causes NetBackup to analyze all options to choose the best one available. Any Available is not recommended.

In general, if the configuration includes many storage units, many volumes within many disk pools, and many media servers, note: the deep analysis that Any Available requires can delay job initiation when many jobs (backup or duplication) are requested during busy periods of the day. Instead, specify a particular storage unit, or narrow NetBackup's search by means of storage unit groups (depending on how storage units and groups are defined).

For more details on Any Available, see the *NetBackup Administrator's Guide, Volume I*.

In addition, note the following about Any Available:

- For Any Available, NetBackup operates in prioritized mode, as described in the next section. NetBackup selects the first available storage unit in the order in which they were originally defined.
- Do not specify Any Available for multiple copies (Inline Copy) from a backup or from any method of duplication. The methods of duplication include Vault, staging disk storage units, lifecycle policies, or manual duplication through the Administration Console or command line. Instead, specify a particular storage unit.

Performance considerations for the Storage Unit Groups method

Although Storage Unit is the preferred method for most large environments, a Storage Unit Group may be useful. It contains a specific list of storage units for NetBackup to choose from. Only these storage units are candidates for the job.

You can configure a storage unit group to choose a storage unit in any of the following ways:

- Prioritized
  Choose the first storage unit in the list that is not busy, down, or out of media.
- Failover
  Choose the first storage unit in the list that is not down or out of media.
- Round robin
  Choose the storage unit that is the least recently selected.
- Media server load balancing
  NetBackup avoids sending jobs to busy media servers. This option is not available for the storage unit groups that contain a BasicDisk storage unit.

You can use the New or Change Storage Unit Group dialog in the NetBackup Administration Console to make the desired modifications. NetBackup gives preference to a storage unit that a local media server can access. For more information, see the NetBackup online Help for storage unit groups, and the *NetBackup Administrator's Guide, Volume I*.

Note: Regarding storage unit groups: the more narrowly defined your storage units and storage unit groups, the sooner NetBackup can select a resource to start a job.

In complex environments with large numbers of jobs required, the following are good choices:

- Fewer storage units per storage unit group.
- Fewer media servers per storage unit. In the storage unit, avoid Any Available media server when drives are shared among multiple media servers.
- Fewer disk volumes in a disk pool.
- Fewer concurrent jobs. For example, less multiplexing, or fewer tape drives in each storage unit.

See "NetBackup job delays" on page 28.

About file system capacity and NetBackup performance

Ample file system space must exist for NetBackup to record its logging entries or catalog entries on each primary server, media server, and client. If logging or catalog entries exhaust available file system space, NetBackup ceases to function.

Veritas recommends the following:

- You should be able to increase the size of the file system through volume management.
- The disk that contains the NetBackup primary catalog should be protected with mirroring or RAID hardware or software technology.

As AdvancedDisk Pools and media server deduplication disk pools get close to full, NetBackup begins limiting the number of concurrent jobs to 1. See the *NetBackup AdvancedDisk Storage Solutions Guide* for more information.

About the primary server NetBackup catalog

The primary server NetBackup catalog resides on the disk of the NetBackup primary server.

The catalog consists of the following parts:

Image database	The image database contains information about what has been backed up. It is by far the largest part of the catalog.
NetBackup data in relational databases	This data includes the media and volume data describing media usage and volume information that is used during the backups.

NetBackup configuration files	Policy, schedule, and other flat files that are used by NetBackup.

For more information on the catalog, refer to "Protecting the NetBackup catalog" in the *NetBackup Administrator's Guide, Volume 1*.

Guidelines for managing the primary server NetBackup catalog

Consider the following:

- Back up the catalog.
  Catalog backup can be performed while regular backup activity takes place. It is a policy-based backup. It also allows for incremental backups, which can significantly reduce catalog backup times for large catalogs.

  Warning: Failure to backup the primary server NetBackup catalog may result in data loss if a catastrophic failure occurs to the file systems housing the various parts of the catalog.

  Note: Veritas recommends schedule-based, incremental catalog backups with periodic full backups.

  Be cautious in using Accelerator full backups daily as a replacement for daily incremental backups. While Accelerator full backups are quick to run, the catalog size will be a full catalog backup instead of an incremental and can grow quickly in size. Backups of client data that contain millions of small files in combination with the use of Accelerator and frequent full backups can also cause the catalog to bloat.

- Store the catalog on a separate file system.
  The primary server NetBackup catalog can grow quickly depending on backup frequency, retention periods, and the number of files being backed up. With the catalog data on its own file system, catalog growth does not affect other disk resources, root file systems, or the operating system.
  Information is available on how to move the catalog.
  See “Methods for managing the catalog size” on page 38.
  See “How to calculate the size of your NetBackup image database” on page 15.
  The following directories and files that are related to the catalog can also be moved. Using an SSD device also improves performance:

On a Linux/UNIX host:

- `/usr/openv/netbackup/db/error` (directory)
- `/usr/openv/netbackup/db/images` (directory)
- `/usr/openv/netbackup/db/jobs` (directory)
- `/usr/openv/netbackup/db/rb.db` (file)

On a Windows host:

- `C:\Program Files\VERITAS\NetBackup\db\error` (directory)
- `C:\Program Files\VERITAS\NetBackup\db\images` (directory)
- `C:\Program Files\VERITAS\NetBackup\db\jobs` (directory)
- `C:\Program Files\VERITAS\NetBackup\db\rb.db` (file)

- Change the location of the NetBackup relational database files.
  The location of the NetBackup database files can be changed for better performance. For example, you may want to change the database location if the default location is running short of space. Using an SSD device also can improve performance.
  The following directories and files that are related to the catalog can also be moved:

  On a Linux/UNIX host:

  - `/usr/openv/tmp` (directory)
  - `/usr/openv/var` (directory)
  - `/usr/openv/db/data` (directory)
  - `/usr/openv/db/staging` (directory)

  On a Windows host:

  - `C:\Program Files\VERITAS\NetBackup\Temp` (directory)
  - `C:\Program Files\VERITAS\NetBackup\var` (directory)
  - `C:\Program Files\VERITAS\NetBackupDB\data` (directory)
  - `C:\Program Files\VERITAS\NetBackupDB\staging` (directory)

  Refer to the procedure in the section *Moving a database after installation* in the *NetBackup Administrator's Guide, Volume I*.

- Set a delay to compress the catalog.
  The default value for this parameter is 0, which means that NetBackup does not compress the catalog. As your catalog increases in size, you may want to use a value between 10 days and 30 days for this parameter. When you restore

old backups, NetBackup automatically uncompresses the files as needed, with minimal performance effect.
See "Methods for managing the catalog size" on page 38.

- Adjust the batch size for sending metadata to the catalog.
  This setting affects overall backup performance, not the performance of catalog backups.
  See "Adjusting the batch size for sending metadata to the NetBackup catalog" on page 37.
- Best practices for primary server NetBackup catalog layout:
  https://www.veritas.com/content/support/en_US/article.100003918

Adjusting the batch size for sending metadata to the NetBackup catalog

You can change the batch size that is used to send metadata to the NetBackup catalog during backups. You can also change the batch size for sending metadata to the catalog specifically for catalog backups.

A change to the batch size can help in the following cases:

- If backups fail because a query to add files to the catalog takes more than 10 minutes to complete.
  In this case, the backup job fails, and the `bpbrm` log contains a message that indicates a failed attempt to add files to the catalog. Note that the `bpdbm` log does not contain a similar message.
- To reduce the CPU usage on primary server.
  When many files are to be backed up and the batch size is small, too many `bpdbm` could be running and using up CPU cycles of the primary server. CPU utilization could be reduced much by increasing the batch size. For large numbers of small file backups, it is recommended to set the value to at least 90000. For NetBackup Flex Scale (NBFS), it is set to 100000 by default.
- To improve backup performance when the folders to back up contain a large number of small files or subfolders.

To adjust the batch size for sending metadata to the catalog for NBU-Catalog backups

1. Create the following file:

   `/usr/openv/netbackup/CAT_BU_MAX_FILES_PER_ADD`

2. In the file, enter a value for the number of metadata entries to be sent to the catalog in each batch, for catalog backups. The allowed values are from 1 to 100,000.

   The default is the maximum of 100,000 entries per batch. Veritas recommends that you experiment with a lower value to achieve the best performance for your backup.

Methods for managing the catalog size

To manage the catalog size, consider the following:

- Why are long-running catalog backups an issue?
- Leverage Differential/Incremental Backups
- Enable Catalog Compression
- In general, large catalogs are the result of long-term retention (LTR) requirements or data sets with large numbers of files. (Typically, NAS filers can have millions of files.) The combination of these two situations can increase the catalog size requirements significantly. Preplanning and creating multiple NetBackup domains in such situations may be an option in very large environments.

  However, defining a domain based upon retention is not a common practice. Clients with large NetBackup environments often plan their domains based upon the workload type groupings rather than upon LTR. Additionally, more backups with LTR are being directed to Access or Cloud S3.

NetBackup has no hard limit on catalog size. However, Veritas recommends as a best practice that you keep the catalog size under 4 TB to ensure good catalog backup and recovery performance. Depending on the size of the environment and the length of backup image retention, catalogs may grow in excess of 4 TB. This size is not an issue for NetBackup, but it can result in operational issues for the environment regarding the time it takes to perform catalog backups and recoveries. This directly impacts the ability to recover the environment in the event of a disaster.

Several methods can be implemented to help mitigate this issue:

- Move the catalog to flash storage.

- Implement incremental or differential schedules for daily backups. This approach can reduce the time required to perform these backups. It can however negatively affect catalog recovery times, and regular full backups are still recommended.
- Implement database compression. This method shrinks the size of the catalog and improves performance of catalog backups.
- Implement catalog archiving. This method shrinks the size of the active catalog, however, it can increase the time that is required to perform restores from archived images.
- Create a separate NetBackup domain for long-term retention. In many cases, excessive catalog size is a result of long-term retention of backup images.

Migrate the catalog

If catalog backups do not complete within the desired backup window, consider moving the catalog to higher performance storage. This method most directly improves catalog backup and recovery performance.

Catalog schedules

Daily differential or incremental backups can be used to ensure that regular catalog protection can be completed within the desired window. For more information on catalog schedules, refer to the *NetBackup Administrator's Guide, Volume I*.

Catalog archiving

If catalog backups do not complete within the desired backup window, consider the use of catalog archiving.

Catalog archiving reduces the size of online catalog data by relocating the large catalog files to secondary storage. NetBackup administration continues to require regularly scheduled catalog backups, but without the large amount of catalog data, the backups are faster.

For more information on archiving the catalog, refer to the *NetBackup Administrator's Guide, Volume I*.

Image database compression

When the image database portion of the catalog becomes too large for the available disk space, you can do either of the following:

- Compress the image database.
- Move the image database.

For details, refer to "About image catalog compression" " and "Moving the image catalog in the *NetBackup Administrator’s Guide, Volume I*.

Note that NetBackup compresses the image database after each backup session, regardless of whether any backups were successful. The compression happens immediately before the execution of the `session_notify` script and the backup of the image database. The actual backup session is extended until compression is complete. Compression is CPU-intensive. Make sure that primary server has enough free CPU cycles to handle the compression.

Long-term retention domain

Long-term retention (LTR) of backup data for multiple years can result in very large catalogs. One method to reduce the effect of LTR is to use NetBackup replication and perform LTR in a separate dedicate domain. This method has the advantage of keeping the catalog for the primary production domain more manageable in size. The catalog in the LTR domain can still become large, however, this problem is of lesser operational effect because rapid recovery of LTR data is generally not required.

Performance guidelines for NetBackup policies

The following policy items may have performance implications.

Table 2-2

Policy items	Guidelines
Include and exclude lists	Consider the following: ■ Do not use excessive wild cards in file lists. When wildcards are used, NetBackup compares every file name against the wild cards. The result may be a decrease in NetBackup performance. Instead of placing `/tmp/*` (Linux/UNIX) or `C:\Temp\*` (Windows) in an include or exclude list, use `/tmp/` or `C:\Temp`. The inappropriate use of wildcards can also flood the policy execution manager (`nbpem`) with thousands of backup jobs for a single client. This situation causes delays in all job processing as `nbpem` takes the necessary actions to start a backup job. ■ Use exclude lists to exclude large unwanted files. Reduce the size of your backups by using exclude lists for the files your installation does not need to preserve. For instance, you may decide to exclude temporary files. Use absolute paths for your exclude list entries, so that valuable files are not inadvertently excluded. Before adding files to the exclude list, confirm with the affected users that their files can be safely excluded. Should disaster (or user error) strike, not being able to recover files costs much more than backing up extra data. When a policy specifies that all local drives be backed up (ALL_LOCAL_DRIVES), `nbpem` initiates request with `nbjm` to connect to the client and runs `bpmount -i` to get a list of mount points. Then `nbpem` initiates a job with its own unique job identification number for each mount point. Next the client `bpbkar` starts a stream for each job. Only then does NetBackup read the exclude list. When the entire job is excluded, `bpbkar` exits with status 0, stating that it sent zero of zero files to back up. The resulting image files are treated the same as the images from any other successful backup. The images expire in the normal fashion when the image header files' expiration date specifies they are to expire. ■ Use exclude lists to exclude files from regular backups if the files are already backed up by a NetBackup database agent backup.
Critical policies	For catalog backups, identify the policies that are crucial to recovering your site in the event of a disaster. For more information on catalog backup and critical policies, refer to the *NetBackup Administrator's Guide, Volume I*. See "Guidelines for managing the primary server NetBackup catalog" on page 35.
Schedule frequency	Minimize how often you back up the files that have not changed, and minimize your consumption of bandwidth, media, and other resources. To do so, limit full backups to monthly or quarterly, followed by weekly cumulative-incremental backups and daily incremental backups.

Legacy error log fields

This section describes the fields in the legacy log files that are written to the following locations:

Linux/UNIX:

```
/usr/openv/netbackup/db/error
```

Windows:

```
install_path\NetBackup\db\error
```

On Linux/UNIX, there is a link to the most current file in the error directory. The link is called `daily_messages.log`.

The information in these logs provides the basis for the NetBackup ALL LOG ENTRIES report. For more information on legacy logging and unified logging (VxUL), refer to the *NetBackup Troubleshooting Guide*.

Here is a sample message from an error log:

```
1021419793 1 2 4 nabob 0 0 0 *NULL* bpjobd TERMINATED bpjobd
```

Table 2-3 defines the various fields in this message. The fields are delimited by blanks.

Table 2-3 Meaning of daily_messages log fields

Field	Definition	Value
1	Time this event occurred (ctime)	1021419793 (= number of seconds since 1970)
2	Error database entry version	1
3	Type of message	2
4	Severity of error: 1: Unknown 2: Debug 4: Informational 8: Warning 16: Error 32: Critical	4
5	Server on which error was reported	`nabob`
6	Job ID (included if pertinent to the log entry)	0
7	(optional entry)	0
8	(optional entry)	0

Table 2-3 Meaning of daily_messages log fields *(continued)*

Field	Definition	Value
9	Client on which error occurred, if applicable. Otherwise *NULL*	*NULL*
10	Process which generated the error message	`bpjobd`
11	Text of error message	TERMINATED bpjobd

Table 2-4 lists the values for the message type, which is the third field in the log message.

Table 2-4 Message types

Type Value	Definition of this message type
1	Unknown
2	General
4	Backup
8	Archive
16	Retrieve
32	Security
64	Backup status
128	Media device

Chapter 3

Media server configuration guidelines

This chapter includes the following topics:

NetBackup hardware design and tuning considerations

This topic contains discussions about the following NetBackup hardware design and tuning considerations:

- PCI architecture
- Central processing unit (CPU) trends
- Storage trends
- Conclusions

PCI architecture

Peripheral Component Interconnect (PCI), PCI-X architecture was the first step of sending PCI signals quickly to the peripheral cards such as Ethernet NICs, Fibre Channel and Parallel SCSI Host Bus Adapters as well as RAID controllers, all of

which enabled RAID storage and advanced connectivity to many servers in a network.

PCI-X was a very good start, beginning in 1998, to the solution. It was a parallel interface and utilized an expander to derive multiple “slots” from the signals sent from the CPUs. With parallel architecture, timing of the signals needed to be rigidly enforced as they all needed to arrive or be sent at the concurrent times. With this restriction, the overall speed and latency of the system was limited to the frequency of the timing circuitry in the hardware. As the speed market needs kept increasing, the difficulty of maintaining the concurrent timing became more and more difficult.

PCIe came into being in 2002 and provided the change to the Peripheral Component Interconnect with two features; serial communication and direct communication from the processor to the PCIe enabled card, NIC, HBA, RAID, and so on. This allowed for a significant increase in bandwidth as multiple lanes of PCIe could be allocated to the cards. As an example, Fibre Channel Host Bus Adapters in 1998 had speeds of 1 Gb with PCI-X, and today, 22 years later, the standard is 16Gb and 32Gb is expected to surpass 16Gb in the next two years.

PCI-X @ 133Mhz was the last widely supported speed. PCIe supplanted PCI-X at 800MB/s data transfer speed. PCIe 3 can today achieve up to 15.745GB/s with 16 lane cards. PCIe 4, which is available today on AMD processor systems and will be available on Intel based systems in 2021, can reach 15.745GB/s with 8 lane cards as PCIe 4 doubles the transfer rates of the current PICe 3 Architecture. The following page notes the speed capability of the versions past and future. By 2026, the supported PCIe throughput is expected to increase 8-fold.

It is expected that the number of PCIe lanes per processor will increase rapidly in the future. It’s easy to see that the race to increase lanes dramatically is on by reviewing currently available processors. For example, the Intel processor family has 40 PCIe lanes, and AMD has countered with 128 lanes per processor.

Table 3-1 PCI Express Link performance

Version	Introduced	Line code	Transfer rate	Throughput				
				1 lane	2 lane	4 lane	8 lane	16 lane
1.0	2003	8b/10b	2.5 GT/s	0.250GB/s	0.500GB/s	01.00GB/s	2.00GB/s	4.00GB/s
2.0	2007	8b/10b	5.0 GT/s	0.500GB/s	1.00GB/s	2.00GB/s	4.00GB/s	8.00GB/s
3.0	2010	128b/130b	8.0 GT/s	0.985GB/s	01.969GB/s	3.938GB/s	7.877GB/s	15.754GB/s
4.0	2017 (now on AMD)	128b/130b	16.0GT/s	1.969GB/s	3.938GB/s	7.877GB/s	15.754GB/s	31.508GB/s

Table 3-1 PCI Express Link performance *(continued)*

Version	Introduced	Line code	Transfer rate	Throughput				
				1 lane	2 lane	4 lane	8 lane	16 lane
5.0	2019 (projected 2022)	128b/130b	32.0GT/s	3.938GB/s	7.877GB/s	15.754GB/s	31.508GB/s	63.015GB/s
6.0	2021 (projected 2024)	128b/130b+ PAM-4+ECC	64.0GT/s	7.877GB/s	15.754GB/s	31.508GB/s	63.015GB/s	126.031GB/s

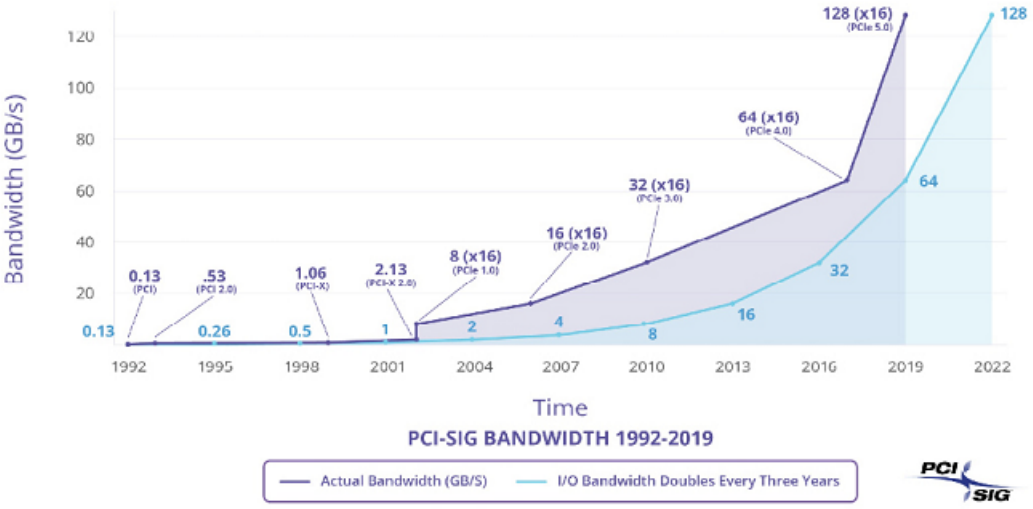
With the advance in speed of the CPU to peripheral communication, both latency (decreased lag) and data transfer rate (more data per second) have improved dramatically. Intel CPUs will increase from present Skylake and Cascade Lake families of 40 PCIe 3 lanes to Ice Lake with 64 PCIe 4 lanes. As noted earlier, AMD had built their processors with 128 PCIe4 lanes. The reason for this up trend is peripherals other than Ethernet, Fibre Channel, and RAID are quickly earning a place on the bus.

NVMe SSDs (Non-Volatile Memory express, Solid State Drives) have quickly carved out a significant niche in the market. The primary advantage they possess is the use of PCIe connection to the processor. These SSDs do not require a SAS or SATA interface to communicate, which results in significant speed and latency advantages because the inherent media conversion is not needed. With the aforementioned PCIe 4 coming into being and the expansion of the number of PCIe lanes, the speed of the NVMe SSD will double, increasing throughput and decreasing (slightly but measurable) access timing.

The latest designs of Intel and AMD motherboards accommodate the NVMe architecture as the primary storage for the future. It is expected that in 2021, systems with the new architecture will be available and will have density up to 12.8TB, speeds of 8,000 MB/s reads and 3,800 MB/s writes and 24 SSDs. These new systems will be dramatically faster than earlier disk-based solutions, which can struggle to reach 10 – 12GB/s reads or writes. The new architecture will also increase network reads and writes; 200GB/s is not a difficult to reach read nor is a 100 GB/s Write. As a scale perspective a 30TB backup at 0% deduplication would take 8 seconds with the proper transport: 304 connections of 100Gb Ethernet NIC ports. Now, this is not going to be the kind of network bandwidth we can expect, but it is illustrative of the coming speeds.

The future of the PCI-e technology is to move to PCI-e 5.0 starting in 2022 and 6.0 in 2024. This cadence would appear to be rather optimistic given the past history of the PCIe revisions as shown below.

Figure 3-1 PCI Bandwidth over time



It should be noted, however, that the specifications for revisions 5.0 and 6.0 are well defined. The significant challenge appears to be, from the delays that were incurred on the 4.0 release, with routing on motherboards. It stands to reason that the PCIe 5 and 6 will be relegated to very high-end systems initially, such as 4 and 8 socket systems that can more adequately use the additional bandwidth and number of lanes.

Central processing unit (CPU) trends

Processors and their relative speed and number of cores have a distinct performance impact on overall system performance. As an example, the latest incarnation of the Intel processors has incremented the number of cores in one of the most popular processors from 20 to 26 per CPU (2 included per motherboard). Through testing it has shown 30% performance improvement at high deduplication rates and a large number of streams. There is also an increase in the speed of the Ethernet NICs from 10 to 25Gb/s, which is the newest “standard” of Enterprise Ethernet. The combination of the two are the only things that have changed from the 5340 system, so the conclusion of additional cores and network bandwidth support the contention of size and speed do matter.

Current processor development is undergoing a significant level of change. AMD has re-emerged and is out pacing Intel with regard to core count and process. At the time of the writing of this document, AMD has a 64 core processor built on 7 nanometer manufacturing process. When compared to Intel at 14 nanometer and maximum core count of 26 it is apparent that Intel is trailing but fighting mightily to

maintain their market dominance. There is a significant movement to create processors utilizing the RISK-V architecture. (See https://riscv.org.) This particular architecture, with the combination of a small number of cores and a collection of coprocessor type cores, is an interesting disruptive technology. It is the author's opinion that this space should be monitored as RSIC-V processors have now been designed and implemented on Apple computers as the first step in moving away from Intel processors.

Figure 3-2 Intel XEON Processor core count and release dates

	1 or 2 Sockets			4 or 8 Sockets		
	3000/5000/E3/E5-1xxx and 2xxx/E7-2xxx series			7000/E5-4xxx/E7-4xxx and 8xxx series		
Node	Code named	# of Cores	Release Date	Code named	# of Cores	Release Date
250 nm				Drake	1	Jun 98
				Tanner	1	Mar 99
180 nm				Cascades	1	Oct 99
	Foster	1	May 01	Foster MP	1	Mar 02
130 nm	Prestonia	1	Feb 02			
	Gallatin	1	Mar 03	Gallatin MP	1	Nov 02
90 nm	Nocona	1	Jun 04			
	Irwindale	1	Feb 05	Cranford	1	Mar 05
				Potomac	1	Mar 05
	Paxville	2	Oct 05	Paxville MP	2	Dec 05
65 nm	Dempsey	2	May 06			
	Sossaman	2	Mar 06			
	Woodcrest	2	Jun 06	Tulsa	2	Aug 06
	Conroe	2	Oct 06			
	Clovertown	4	Nov 06			
	Allendale	2	Jan 07			
	Kentsfield	4	Jan 07	Tigerton	2	Sep 07
45 nm	Wolfdale DP	2	Nov 07			
	Harpertown	3	Nov 07			
	Wolfdale	2	Feb 08			
	Yorkfield	4	Mar 08	Dunnington	4-6	Sep 08
	Nehalem-EP	2/4	Mar 09			
	Bloomfield	4	Mar 09			
	Gainestown	2/4	Mar 09			
	Beckton (65xx)	4/6/8	Mar 10	Beckton (75xx)	4-8	Mar 10
32 nm	Westmere-EP (56xx)	2-6	Mar 10			
	Gulftown (W36xx)	6	Mar 10			
	Westmere-EX (E7-2xxx)	6-10	Apr 11	Westmere-EX (E7-4xxx/8xxx)	6-10	Apr 11
	Sandy Bridge-EP	2-8	Mar12	Sandy Bridge-EP (E5-46xx)	4-8	May 12
22 nm	Ivy Bridge (E3/E5-1xxx/E5-2xxx v2)	2-12	Sep 13	Ivy Bridge-EP (E5-46xx v2)	4-12	Mar 14
	Ivy Bridge-EX (E7-28xx v2)	12/15	Feb 14	Ivy Bridge-EX (E7-48xx/88xx v2)	6-12/15	Feb 14
	Haswell (E3/E5-1xxx/E5-2xxx v3)	2-8	Sep 14	Haswell-EP (E5-46xx v3)	6-18	Jun 15
				Haswell-EX (E7-48xx/88xx v3)	4-18	May 15
14 nm	Broadwell (E3/E5-1xxx/E5-2xxx v4)	4-22	Jun 15			
	Skylake-DT (E3 v5)	4	Oct 15			
	Kaby Lake-DT	4	Mar 17			
	Skylake-X	6-18	Jun 17	Skylake-SP	4-28	Jul 17
	Cascade Lake-X	10-18	Nov 19	Cascade Lake-SP	4-28	Apr 19
				Cooper Lake-SP	16-28	Jun 20

The conclusion drawn from the Cores / PCIe sections above is that significant changes are at our doorstep now and we need to be experimenting with the various platforms to find the best combination. Note the tables above as they detail the evolution of the Xeon processor, especially the number of cores. When NetBackup 7 debuted in 2010, the largest number of cores in a 2U system was 12 using two 6 core processors. 11 years later 52 cores are a popular configuration, an increase of 4 cores per year. This "trend" is not going to change in the near future and as

such, when building a system, users must allocate enough cores within the processor(s) to address the number of concurrent operations.

Storage trends

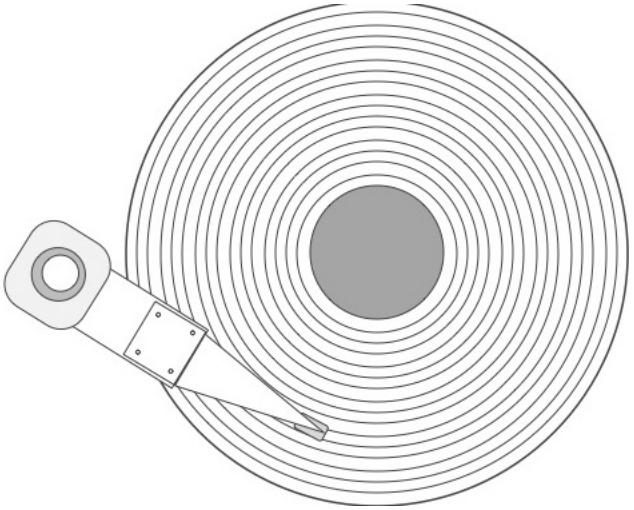
Solid state disks (SSDs) as storage on our systems, NetBackup + Customer Platform, Veritas Appliances and Reference Architecture + NetBackup will be the norm rather than the exception in the near future.

A chart from BlockandFiles.com provides a good indicator of the possible scenario we will encounter. SSD is in its early phase while disk is in its late stages of life. We can look at other technologies that were superseded. A relevant example is disk versus tape. Tape has historically been a excellent long term retention media and disk a short-term retention device. With the ongoing increase in data creation, the size of it necessitates a new type of storage. We will likely see the change in the next few years where SSDs push disk (including cloud providers) to long term retention and tape relegated to retention that is legislate mandated such as 30 year for medical information.

Disk drive companies have been trying to make a heat or microwave assisted magnetic recording, known as HAMR or MAMOR disk drives for more than 20 years. This technology portends to increase the bit density on the drive platters to provide storage in the 20 to 50TB per drive range. However, there is yet to be a viable example of a drive with those capacities at a cost that significantly drive the $/GB down past the $0.02 level. Meanwhile the SSD is continuing its increase in performance and decline in cost per GB.

The key advantage of SSDs versus disk for the NetBackup software is the access time advantage SSDs possess. Current SAS Nearline Disk Drives are specified at as low a 4.16 millisecond average rotational latency and 8 millisecond seek for a total of 12.16 millisecond access time. NVMe SSD are in the 10-microsecond range at this time. SSDs do not suffer from some of the mechanical requirements of disk drives.

Figure 3-3 Disk seek time



Drives require that the Read/Write head be positioned properly over the sector, which resides in a track, before it can be read or written to. This is not a consistent exact operation as the tracks have an extremely small width to them. To get a perspective, the drives used presently have a track density of 371,000 tracks per inch. Each of the platters of the drive have 337,400 tracks per disk surface. This works out to 0.0000027 inch width per track. Because of this, there is not always a perfect “settle” over the sector and the disk will retry. This requires a full revolution of the disk which results in an 8.33 millisecond latency for each retry.

Since disk drives are a mechanical device, they are subject to external forces, especially vibration. Drives have adopted Rotational Vibration Accelerometers (RVA) to compensate for the external vibrations, but they are not effective on all impetus. Vibrations from adjacent drives or other external sources can have enough amplitude to overwhelm the RVA and cause a retry. Retries can be numerous and retries of up to 20 times happen. If the retries exceed a 60 second window, the RAID controller will enact a SCSI Time Out, the drive will be marked as bad, and removed from the RAID set.

When drives fail, they are typically returned to the manufacturer for failure analysis. 80% of failed drives returned are found to be No Trouble Found (NTF) and the reason is usually from a SCSI Time Out as noted previously. Vibration and track settle are the most likely cause of this failure and as such, creates additional cost, performance degradation during rebuild, and lost time.

As the market has evolved, access time has become more and more important. Sequential time in the era of tape was vital as it would allow for more data per cartridge with higher throughput. Now, with deduplication, the speed to read from

the client is very important, but in the vast majority of backup environments, the need for throughput to the target disks is diminished. The key in the present era, in addition to up time, is Access time.

Conclusions

When specifying and building systems, understanding the use case is imperative. The following are recommended courses of action depending on the use case.

Processors

The large number of concurrent streams needed for nightly backups requires higher number of cores per processor. If looking at an enterprise-level backup it is recommended that 40 to 60 cores per compute node are required. More is not necessarily better, but if the user is backing up very large numbers of highly deduplicatable files, a high number of cores are required.

Mid-range stream requirements indicate a 12 to 36 core system. This assumes that the requirements are approximately 20 to 70% of the workload of the enterprise environment as shown above.

Small systems should look at 8 to 18 core systems and single processor motherboards as they will reduce cost and accommodate today's processor core count.

DRAM memory

Quality dynamic RAM (DRAM) is extremely important to ensure accurate operation. Because of the number of concurrent backups that users look to accomplish, Error Code Correction (ECC) and Registered (R) DRAM are required to ensure operation with no issue. Current systems use DDR4 SDRAM as the abbreviated "Double Data Rate Synchronous Dynamic Random-Access Memory" with the 4 representing the fourth generation of DDR memory. Users must use DDR4 ECC RDIMMs with current, as of the writing of this document, processors. Frequencies and generation of the DRAM must align with the processor recommendation and be of the same manufacturing lot to ensure smooth operation.

Current requirements of RAM in backup solutions are tied to the amount of MSDP data that is stored on the solution. To ensure proper and performant operation, 1 GB of RAM for every terabyte of MSDP data is recommended. For instance a system with 96TB of MSDP capacity requires the use of at least 96GB of RAM. DDR4 ECC RDIMMs come in 8, 16, 32, 64 and 128GB capacity. For this example, 12 each 8GB DIMMs would suffice, but may not be the most cost effective. Production amounts of the different sizes will change the cost per GB and the user may find that a population of 6 each 16GB or even 8 each 16GB, 128GB total may

be a more cost effective solution and provide a future path to larger MSDP pools as the need for such increases.

PCIe

When selecting a system or motherboard, it is recommended that a PCIe 4 compliant system be chosen. In addition to the doubling of speed of the PCIe lanes, the number of lanes on processors will increase thereby creating a more than 2X performance enhancement. PCIe 4 Ethernet NICs, up to 200Gb, Fibre Channel HBAs up to 32Gb, SAS HBAs and RAID controllers at 4x10Gb per port all with up to 4 port or port groups can take advantage of this higher bandwidth. This level of system will be applicable for 7 to 10 years as opposed to PCIe 3 level systems that will likely disappear in the 2023 time frame. Users will be able to continue to utilize PCIe 3 based components as PCIe 4 is rearward compliant. However, it appears that the PCIe 4 components are in the same price range as PCIe 3, so the user is encouraged to utilize the newer protocol.

Disk drives and RAID storage

Disk drives have the potential to have rather large capacity in the future. HAMR and MAMR as noted earlier are technologies poised to create large, petabyte to exabyte scale repositories with up to 50TB drives. Assuming that consumption continues a 30% per year expansion, these sizes will fulfill the needs of backup storage for the foreseeable future.

For build-your-own (BYO) systems with present day 256TB capacity the best solution would be to design storage that brackets the 32 TiB volumes. For instance, using RAID 6 volumes with a hot spare, as the Veritas NetBackup and Flex appliances use, it is wise to create volumes that can contain those sizes of volumes efficiently. As an example, the NetBackup and Flex 5250 appliances utilize a 12 drive JBOD connected to a RAID controller in the Main Node. It uses 8TB drives and with a RAID 6 using 11 of the drives +1 for hot spare the resultant capacity is 72TB / 65.5 TiB. With this, two volumes of 32TiB fit well into the JBOD and can easily be stacked to arrive at the maximum capacity.

Solid-state disks (SSDs)

SSDs present a new variable into the solution as they act like disk drives but are not mechanical devices. They present lower power, high capacity, smaller size, significant access time improvement over disk and higher field reliability. The one downside, as compared to disks, is cost. For certain implementations though they are the best solution. Customers who require speed are finding that SSDs used for tape out of deduplicated data are 2.7 times faster than disk storage. If concurrent operations are required such as backup and then immediate replication to off-site, the access time of the SSDs used as the initial target make this possible in the time window necessary. Another use case is to use the SSDs as an Advance Disk pool

and then, after the user feels the time is appropriate, the data could be deduplicated to a disk pool for medium or long-term retention.

As noted, earlier NVMe should be the choice for the best performance. Expectations are that the Whitley version of the Intel reference design, due for release in 2021, will be the best Intel platform as it will feature PCIe 4. With the incremental doubling of speed, only 2 lanes would be necessary allowing for an architecture that can handle a large number of SSDs, 24 in a 2u Chassis as well as accommodate the requisite Ethernet and Fibre Channel NIC/HBA to connect to clients.

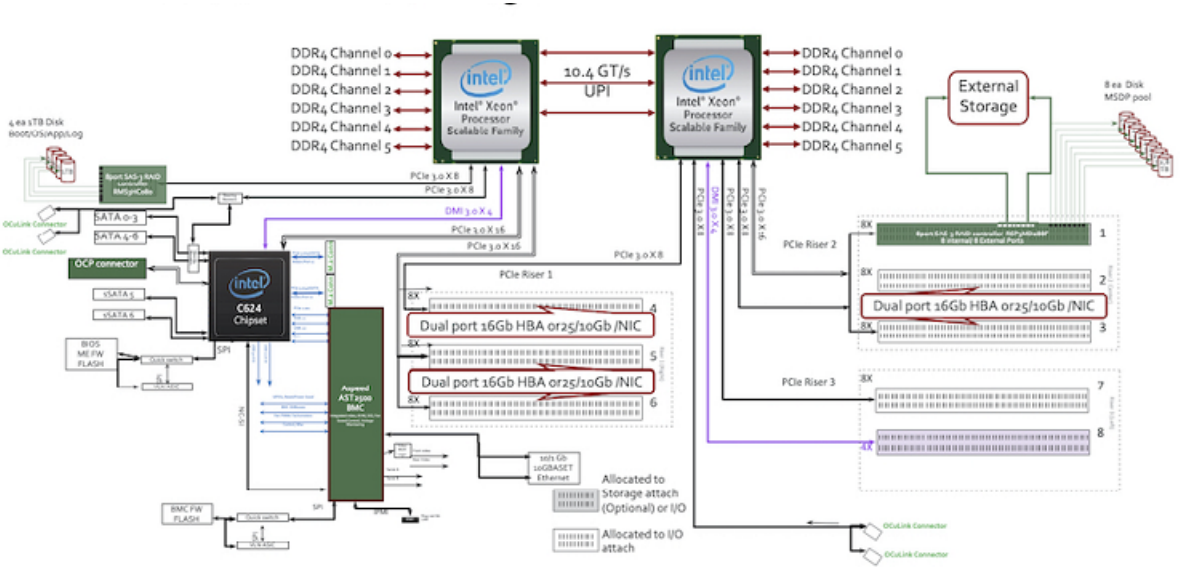
Ethernet

As the predominant transport for backup, Ethernet NICs are of critical importance. Fortunately, there are a number of quality manufacturers of the NICs. For the time being, greater than 90% of the ports used will be 10GBASE-T, 10Gb optical or direct-attached copper (DAC) and 25Gb Optical / DAC. Broadcom and Marvell have NICs that support all three configurations. Intel and NVIDIA have 25-10 Optical / DAC NICs as well as10GBASE-T equipped NICs. Any of these can be used to accommodate the user's particular needs. Forecasts show that 50 and 100 and, to a lesser extent, 200 and 400Gb Ethernet will be growing quickly as the technology advances.

Fibre Channel

Fibre Channel (FC) will continue to exist for the foreseeable future, but much of its differentiation from other transports is lessening as NVMe over fabric becomes a more prevalent solution. FC is one of the transports, but it appears that Ethernet will have the speed advantage and will likely win out as the favored transport. For customers with FC SANs Marvell and Broadcom are the two choices for Host Bus Adapters as initiators and targets. Both are very good initiators, and the choice is up to the user as many sites have settled on a single vendor.

Figure 3-4 Media server block diagram



About NetBackup Media Server Deduplication (MSDP)

The NetBackup Media Server Deduplication (MSDP) feature lets you choose where in the backup process to perform deduplication. In particular, you can choose deduplicating the data at the source (client) or the target (NetBackup server) side. The system resource usage pattern and thus the performance tuning will be different based on the choice.

To facilitate performance planning and tuning, a brief introduction of the technology is included here. For a more detailed description of the technology, refer to the NetBackup Deduplication Guide.

MSDP deduplication technology is composed of four main components:

1. Data segmentation: This component is responsible for dividing a data stream into segments and performing fingerprint calculation using, starting with NetBackup 8.1, SHA-2 algorithm against each segment. With proper data segmentation, you can achieve a higher deduplication ratio when the data stream has interspersed data changes.
2. Fingerprint lookup for deduplication: For each newly created fingerprint, this component compares the new fingerprint against the existing fingerprints that

are cached in memory. Store a pointer but not the corresponding segment if a match is found; if there is no matching fingerprint found in the cache, then the corresponding data segment is written to the storage pool. Storing only unique data segments results in a significant reduction in the storage pool.

3. Data store: This component manages the storing and maintaining the data segments on the persistent storage. It also facilitates the read, write and deletion operations of the stored data segments.
4. Space reclamation: When a data container in the data store which holds a number of data segments is no longer referenced due to delete or image expiration, the space occupied by the container can be reclaimed, and a container having enough deleted space may go through a compaction operation to get the deleted space released. The storage space reclamation is handled by this component to maintain a robust, efficient, and well performing data store.

Data segmentation

MSDP supports two types of segmentation:

1. Application data format-based segmentation: NetBackup provides a number of stream handlers which were specifically developed to process the unique application data streams, such as Oracle, MS SQL Server, and NDMP NAS. Stream handlers break up the data stream into segments which results in higher deduplication by identifying similar data segments within the data stream and across multiple backup images. Stream handlers are also capable of separating image meta data from actual application data resulting in higher deduplication rates. This type of segmentation is less CPU intensive and higher performance than context aware segmentation.
2. Data stream context-aware segmentation, also known as variable length deduplication (VLD): the algorithm does not need to fully understand the data format, but it scans the data to identify anchor points in the data stream by moving a predefined window byte by byte to efficiently generate a series of hash results to identify the anchor points when the hash results match predefined values, and then segments are formed based on the anchors. Window sizes are configurable to allow customized tuning for specific workloads. This type of segmentation is application data agnostic, but usually runs slower as it is more CPU intensive.

Fingerprint lookup for deduplication

The SHA-2 hashing algorithm is used to generate the fingerprints of the data segments from backup streams. A unique SHA-2 fingerprint represents a unique

data segment and is compared to a set of fingerprints representing data segments already in a data store. A lookup match means the data segment is already stored in the system; a lookup miss means the system does not have it and the corresponding data segment needs to be stored.

The set of fingerprints in memory, also known as the fingerprint cache, contains two sets of fingerprints for a given backup job:

1. The global fingerprint cache, which is indexed for fast query, maintained at the deduplication server-side for the duration of the deduplication service running.
2. The job-based fingerprint cache, which is also indexed, created at the deduplication client side in the beginning of the job and released at the end of the job.

   The fingerprints of the last image (which is the last full backup by default and can be the last full backup plus subsequent incrementals) is fetched from the MSDP server to the OST pdplugin in the beginning. Whether the deduplication happens on the OST pdplugin completely depends on whether the client-side cache is big enough to hold all the fingerprints from the last image. Any fingerprint lookup that is missed from the client-side cache triggers the lookup to go to the MSDP server-side, even though the fingerprint may not exist on the server-side.

This two-level fingerprint cache provides a high-performance lookup and reduces memory footprint requirement at the server side, such as a NetBackup appliance.

Predictive and sampling cache scheme

Beginning with NetBackup 10.1, a new fingerprint (FP) cache lookup data scheme was introduced. The new scheme splits the current maximum cache size `MaxCacheSize` into two components, predictive cache (P-cache) and sampling cache (S-cache).

The P-cache is used to cache the fingerprints that are most likely used in the immediate future.

The S-cache is used to cache a percentage of the fingerprint from each backup and a subset of each sample fingerprint is inserted into the S-cache. P-cache is first used to find duplicates, and lookup misses reaching a threshold are searched in S-cache for possible matches. If found, the predicted relevant fingerprints are loaded from disk into the P-cache for deduplication.

For more information about P-cache and S-cache, refer to the *NetBackup Deduplicaton Guide* for 10.1 or later.

With NetBackup 10.1, the P-and-S-cache is the default FP lookup scheme for cloud LSUs (logical storage units), while the local LSU volume is still defaulted to using

the `MaxCacheSize`. The configuration changes and default values for P-and-S-cache cache are listed in the following table:

Table 3-2 Configuration change and default value for P-and-S-cache

Configuration	Default Value
`MaxCacheSize`	512 MB
`MaxPredictiveCacheSize`Max	40% in NetBackup 10.1 10% in NetBackup 10.1.1)
`MaxSamplingCacheSize`	10%
`EnableLocalPredictivesSamplingCache` in `spa.cfg`	true
`EnableLocalPredictiveSamplingCache` in `contentrouter.cfg`	true
`MaxCloudCacheSize`	Deprecated and replaced with Max P-cache size and Max S-cache size

With the above change, to ensure that memory is used for uploading, the formula before NetBackup 10.1 is changed to:

`MaxCacheSize` + `MaxPredictiveCacheSize` + `MaxSamplingCacheSize` + `MaxCloudCacheSize` (Cloud in-memory upload cache size) must be less than or equal to the value of `UsableMemoryLimit`.

With P-and-S-cache in 10.2, local and all cloud LSUs share the same P-and-S-cache, and the previous `MaxCacheSize` can be ignored. The P-and-S-cache setting needs to be done carefully. Setting them too high will waste memory, while setting them too low will lead to a poor deduplication ratio and impact backup performance.

In general, S-cache size should be proportional to the backend storage size, while P-cache size is determined by the maximum number of concurrent jobs. Use the following rules of thumb for the P-and-S-cache tuning:

- For each 10 TB of backend storage, allocate 1 GB of RAM for S-cache
- For each backup stream, allocate 250 MB of RAM for P-cache. So, the total P-cache allocated should be (250 MB) * (maximum number of concurrent jobs)

To ensure enough memory for other processes running on the system, P-and-S-cache size together should not exceed the `MaxUsableMemory` value.

Other processes that also need memory include:

- Basic operating system with NetBackup if running as a media server
- NetBackup processes if NetBackup runs in the same node
- spad cache for the opt-dup source
- mtstrd cache for the backup source
- Spooler cache

Disk cache for cloud upload and download

The NetBackup cloud tier allows each media server to create one or more cloud logical storage units (LSUs). It is important to know that for each cloud LSU created, roughly 1 TB of MSDP storage pool is reserved for the LSU to be used as cloud disk cache.

Starting with NetBackup10.2, this preserved disk cache can be configured from the NetBackup web UI during LSU creation. The disk cache size for upload is 12 GB and is set by the parameter `UploadCacheDB`, while the default disk cache size for cloud download is 1 TB which is set by the parameters `DownloadDataCacheGB` and `DownloadMetaCacheGB`. The default values for parameters are set in `contentrouter.cfg` with `CloudUploadCacheSize`, `CloudDataCacheSize`, and `CloudMetaCacheSize` respectively.

As mentioned earlier, the disk caches occupy space in the MSDP pool. For the MSDP pool with limited storage size, the reserved disk cache can consume too much space, resulting in little usable space for regular backup jobs. If jobs are failing with error codes 129 and 84, it may indicate that there is no space left on the device, even though the MSDP pool may still have plenty of space according to `df -h` and `dsstat`. For this kind of case, we recommend:

- Limit the number of cloud LSUs created per media instance, especially if storage pool is relatively small.
- Reduce the default `CloudDataCacheSize` and `CloudMetaCacheSize` values.

If there is enough memory for upload to go through the memory cache, the `UploadCacheGB` can be set to (maximum number of concurrent streams * `MaxFileSizeMB` * 2) in the `cloud.json` file. If the maximum number of concurrent streams is 100, the `UploadCacheGB` value can be set to 12 GB. The `DownloadDataCacheGB` and `DownloadMetaCacheGB` used for the restore or opt-dup download cache can be as small as a few GBs to function. A larger download disk cache size can improve restore and opt-dup performance because it can help avoid downloading the same data object more than once.

Tuning the `DownloadDataCacheGB` and `DownloadMetaCacheGB` values requires knowing the maximum number of concurrent download streams. In most cases, restoring from the cloud requires downloading the entire data container (64 MB).

This is because the container created at backup time usually consists of data from a single client, and MSDP-C will download the entire container so that the same container is only fetched once during a restore.

The default values of the parameters are set under `<storage>/etc/puredisk/contentrouter.cfg`, and the default values are used for all future LSUs. The parameters in `cloud.json` are used to set values used for each already created LSU. The file is found at `<storage>/etc/puredisk/cloud.json`.

Data store

MSDP stores data segments from a backup data source in an efficient way to minimize fragmentation and to provide sustainable performance over time at scale.

Good data locality is maintained over time through the following techniques:

1. Manage space from multiple file systems created from high performance RAID storage system for high capacity and high parallel file operations. New file systems may be created and added to the pool when additional storage is added.
2. Minimize the number of writes-per-file-system to reduce file system fragmentation and improve space allocation.
3. Create a small percentage of segment duplicates by gathering scattered segments among data containers into a single container to reduce image-level segment fragmentation.
4. Improve restore performance with read-ahead prefetching techniques based on optimized segment read sequences.

Space reclamation

The following storage space reclamation techniques help to maintain a robust, efficient, and well performing data store:

- Manage data reference at data container level which allows for the potential of reclaiming a whole container when the reference to the container drops to 0.
- Maintain statistics on how much space may be reclaimed from a container to trigger compaction operation at the right time on a set of containers for segment level space reclamation.
- Reference count increments are done before a backup job is completed to ensure backup image data is correctly protected.

- Reference count decrements from image expirations are batch processed for efficiency.

System resource usage and tuning considerations

Different backup images achieve different deduplication ratio and can create very different performance bottleneck on the NetBackup server. In general, servers that process many concurrent high deduplication ratio backups, 75% or above, tend to consume high % of CPU and more likely to bottleneck on CPU and network, while low deduplication ratio backups, 60% or below, tend to be I/O bound due to more data segments need to be written to the data store. A clear understanding of your workload and how MSDP consumes the four major system resources, CPU, memory, network and I/O bandwidth is critical for configuring MSDP for optimal performance.

Memory considerations

Beside faster backup, one of the core advantages of MSDP is storage space saving. The amount of storage space saving is determined by the deduplication ratio of the backup images, and the deduplication ratio is affected by the fingerprint lookup cache hit ratio. By default, MSDP is allowed to use up to 50% of the system RAM for fingerprint caching, the cache size for fingerprint is specified by the parameter, MaxCacheSize, defined in the MSDP configuration file, contentrouter.cfg. For older NetBackup releases, 8.2 and lower, the default was set at 75%. Beginning with NetBackup 8.3, the default is changed to 50%.

When the system is new, few fingerprints exist to consume the allocated cache, only a very small amount of the 50% RAM will be used, the system may show lots of free memory in the beginning. However, as more data is backed up to the system, more and more unique fingerprints will be created and more RAM will be consumed to cache these fingerprints. With SHA-2 algorithm, each fingerprint takes 48 bytes of RAM and can't be paged out. The fingerprint cache size and the average segment size determine how many fingerprints would be stored in memory for fingerprint lookup after a client-side cache lookup miss and may impact the deduplication rate. Big backup images, multi-streaming backup images and database multi-channel backup images rely more on the global fingerprint cache for better deduplication since the client-side fingerprint cache is too small to hold the corresponding fingerprints. We recommend configure at least 1GB of RAM for each TB of deduplication storage pool on the MSDP server to achieve a good deduplication rate.

Limiting the MaxCacheSize is necessary to prevent MSDP FP fingerprint caching mechanism from over consuming the RAM and creating the potential memory starvation as more unique fingerprints are created and cached. Note that beside FP cache, the system RAM is also needed for other software components running

on the server, including OS, File System and other NetBackup processes, such as bptm, bpdm, and so on.

In addition, memory is also needed to maintain the following MSDP objects:

- Cloud tier fingerprint cache
- Data stores
- Task related spooler memory
- Memory manager
- Restore prefetch pools
- Compaction operations
- Reference database operations
- Request queues

I/O considerations

MSDP works best with multiple file systems configured to provide the disk space to store its data and metadata. Ideally, each file system should be created on independent storage volumes with equal size and no disk/LUN sharing for best parallel I/O operations. If possible, MSDP metadata should be configured to be stored on a different file system separate from the file systems storing the data containers due to its different I/O patterns.

The size of a file system configured for an MSDP pool may be in the range of several tens of TB up to 100 TB. For each file system, MSDP has dedicated worker threads and data buffers for data writes, compactions, and CRC checks, etc. If the size of the file systems for data containers vary a lot, the smaller file systems may be filled up earlier than the larger ones. The smaller file systems stop receiving data when filled up which reduces the I/O bandwidth and eventually affects server performance. The impact will be more significant, especially for I/O intensive workloads, such as low deduplication ratio backups, optimized duplication, or restores, the last two are read-intensive as the data needed is not likely found in file system cache.

Network considerations

The inbound bandwidth of the aggregated network interface cards determines how fast data may be ingested into the server, including backups, replication and duplication, and so on. The outbound bandwidth determines how fast data may be retrieved from the system, including restore, replication, and duplication, and so on. For a given backup/restore client, the network bandwidth and the connection speed between the client and MSDP determines the maximum throughput achievable between the client and the server. The network latency and package drop rate may

impact the user observed performance as well. For a high latency and low bandwidth network connection, client-side deduplication with multistream daemon (`mtstrmd`) can help boost the backup performance. `mtstrmd` is the MSDP deduplication multithreaded agent. It receives data from the deduplication plug-in through shared memory and processes. It uses multiple threads to improve performance. It performs fingerprint batch query to reduce the performance impacts from the network latency.

See the NetBackup Deduplication Guide for details about `mtstrmd` and how to configure the agent.

CPU considerations

Each running process consumes certain amount of CPU cycles. The main processes consuming CPU cycles include NetBackup processes such as bptm, bpdm, and MSDP spoold, spad, mtstrmd, and vpfsd if Universal Share and Instant Access are used. More concurrent jobs correspond for more bptms, bpdms processes running at the same time, and more active threads in spoold, spad, mtstrmd and vpfsd.

For target deduplication, segmentation, fingerprint calculation and lookups are done on the same host running spoold, spad and vpfsd. More target deduplication jobs use more CPU cycles from the host. When a media server is bottlenecked on CPU, network and/or memory, adding one or more NetBackup load balancing fingerprinting media servers and/or NetBackup Client Direct (client-side deduplication) ease the CPU, memory and network pressure by offloading some data segmentation and fingerprint calculations to the new server.

OS tuning considerations

Due to each customer's unique H/W configuration and workload, each customer's environment may generate different system bottleneck and require tailored OS tunings. NetBackup technical support published known OS tunings that help with system performance and will continue to do so in the future as new tunings are discovered. Beyond those already published tech notes, Veritas recommends running with the default OS setting first. If, as more loads are added, throughput continues improve and the four major system resource usage increase proportionally, then most likely the system is tuned properly. You will know when the system requires tuning, if there is unexpected hardware resource bottleneck or throughput does not improve while plenty of h/w resources is still available. See the flow chart in Chapter 5 for the process of performance issue troubleshooting.

Following are symptoms and remedies for some common system resource bottlenecks that NBU servers may encounter:

Table 3-3

Resource	Symptom	Recommendaiton
Memory	Heavy swapping	**1** Reduce the load. **2** Reduce MaxCacheSize. **3** OS tuning, such as changing the swappiness setting in Linux to force flushing the file system block cache first before swapping out active process data.
Network	Poor network traffic load balancing	**1** Try different bonding algorithm. **2** Add NICs.
	Client connection failure due to exceeding max socket connections	Example: Set net.core.somaxconn=1024 (default=128) on Red Hat OS
CPU	CPU idle dropped to below 15%	Try reducing the load.
I/O	Poor backup performance during Read/write mixed workload	If you use Veritas Infoscale for MSDP storage pool management, the poor performance could be caused by the memory map (mmap) lock contention. For NetBackup versions 8.2, 8.3, 8.3.0.1, 9.0, and 9.0.0.1, contact Veritas support for the EEB that fixed the mmap lock issue.
	Poor I/O performance due to LUN sharing	Avoid LUN sharing by relocating the second volume to a dedicated LUN if one exist
Open files	Maximum open files exhausted	Try doubling the current setting

MSDP tuning considerations

The default MSDP configuration, defined in the `contentrouter.cfg` file should work for most of the installations, in particular for storage pool size 95 TB or less and moderate workload environments.

For large servers that need to support a high number of concurrent jobs (400 or above) and a storage pool larger than 256 TB, some tunings listed in the following table may be necessary.

Table 3-4 MSDP tuning considerations

Parameter	Default value	Recommended value	Purpose
`MaxCacheSize`	75% on older NeBackup versions	50%	Limit the maximum amount of RAM used for fingerprint caching to 50% of the RAM.
`MaxConnections`	3000	8192	Supporting large numbers of concurrent jobs.
`AllocationUnitSize`	2MiB	8MiB	To avoid MSDP engine crashes with stack overflow caused by the process exceeding the maximum number of memory map areas that a process may have. The maximum number is defined by the Linux kernel parameter, `max_map_count`. For systems with RAM size greater than 512 GB, this tuning parameter is highly recommended

Sample steps to change MSDP `contentrouter.cfg`

MaxCacheSize:

```
ssh appadmin@<msdp_instance_name>
sudo /usr/openv/pdde/pdag/bin/pdcfg --write
/mnt/msdp/vol0/etc/puredisk/contentrouter.cfg --section
 CACHE --option MaxCacheSize --value 50%
```

Check the changed MaxCacheSize as below:

```
sudo /usr/openv/pdde/pdag/bin/pdcfg --read
/mnt/msdp/vol0/etc/puredisk/contentrouter.cfg --section
 CACHE --option MaxCacheSize
```

Restart the pdde-storage process with following commands:

```
sudo /etc/init.d/pdde-storage force-stop
sudo /etc/init.d/pdde-storage start
```

MSDP sizing considerations

When considering the number and size of MSDP pools to deploy in a specific NetBackup domain, there are some important pieces of information that should be gathered about the size and type of workloads that require protection. Furthermore, the requirements around secondary operations like replication and/or duplication are also an important consideration, as well as the retention of each operation.

Note: Information about sizing calculations for MSDP is available:

See "Sizing calculations for MSDP clients" on page 22.

The Recovery Point Objective (RPO), the Recovery Time Objective (RTO), and the Service Level Agreement (SLA) also drive the protection policy requirements. Each of these key requirements must be defined as they directly impact the solution design.

Furthermore, designing a solution which includes practicality and resiliency must include factoring in time for routine maintenance and avoiding Single Points of Failure (SPOFs). Expecting a NetBackup domain to never encounter any maintenance is unrealistic. So, any solution design should allow for enough headroom to allow workloads to be shifted if a single pool or hardware component becomes unavailable.

Workload grouping by type is also very important. Different workloads require different compute requirements. For instance, Oracle and VMware workloads tend to be more resource intensive.

Once workload protection requirements are defined, and the size and type of workloads are qualified, determining the sizing and solution design is more straightforward.

Many people don't define solution requirements because they aren't clear about the process and what information is key to making decisions around solution design.

There are very clear variables that must be defined when designing a data protection solution:

- Workload types and sizes
- Data characteristics
- Backup methodology and retentions
- Data lifecycle
- Timing of backup and secondary operations
- RTO, RPO, and SLAs

In the following sections, each step is addressed to clarify the process.

Data gathering

Workloads

Define the workload types that you have and the front-end terabytes (FETB) of each workload. A workload type would be classified as VMware, Oracle, SQL, MS-Exchange, NDMP, etc. For each workload type, then determine if there are some key data characteristics that could be of great importance.

For instance, if Enterprise Vault (EV) is one of the workload types, then a key data characteristic is that it can sometimes result in millions of little files which results in slower, resource intensive backups. Sometimes EV data can be grouped into Windows Cabinet Files (CAB), but if WORM is a factor, then CAB collections aren't an option which means being faced with protecting a workload type with a data characteristic of millions of files with an approximate size of 40KB per file. This is excluding databases and indexes that must be protected as part of a full EV backup.

If the workload types are Oracle and/or SQL, then understanding if transaction and/or archive logs are required to be protected is also very important as it can result in thousands of tiny jobs. The overhead of thousands of tiny jobs versus a smaller number of larger jobs has been observed to be a significant factor in determining compute requirements.

If VMware is the workload type, then often it makes sense to leverage the Accelerator feature to improve performance of VMware backups. The use of that feature leverages additional compute resources, but only for the first copy of the data.

In the case that a specific workload leverages 3rd party encryption, it is recommended that the customer consider leveraging our native encryption instead of a 3rd party encryption. In some cases, a customer may require a specific workload to be encrypted at the source. If this is case, any backups of this data will experience very poor dedupe rates. Therefore, the use of MSDP isn't a good fit for this type of workload. That said, understanding this requirement is very important as it can have a significant impact on solution design, feature use, and solution sizing.

Once the workload qualification is done, calculate how large a complete full backup would be for each workload type. Then, discuss with stakeholders the estimated daily rate of change per workload type. That information is also very important.

Data lifecycle

Complete a data lifecycle for the protection strategy. A data lifecycle traces each step of the primary and secondary processing of a specific type of workload. For example, for a specific workload, is there a secondary operation required? If so, is the secondary operation a duplication, or perhaps a replication via Auto-Image

Replication (AIR)? Also, are there additional steps in the process that involve writing an N+1 copy of the data to S3, or perhaps to tape for offsite storage? Tracing each step is critical because each step will be part of a NetBackup Storage Lifecycle Policy (SLP), which then requires resources to complete.

Backup methodology and retention

It is important to consider what the retention requirements are for each copy. Perhaps the primary copy of the data would be retained in an MSDP pool for 35 days and the secondary copy via AIR is kept for 90 days. Then, perhaps this workload requires a long-term retention (LTR) whereby a 3rd copy is sent to an S3 bucket or out to tape for a period of several years.

As part of this step in defining the retention requirements, determine what type of backup is required and how often, as well as which type of backups are targeted for LTR. Is it a weekly full? A monthly full? Typically, frequent backups being duplicated or replicated to LTR aren't a good fit because of cost and long-term management of such a large amount of data. Consider the S3 storage costs for storing large amounts of data with an LTR of many years. For customers that require a tape copy, consider the management implications of an LTR of potentially hundreds of tapes offsite for many years. There is a cost of time and compute resources to produce a tape copy, and then the cost of maintaining that data in the NetBackup catalog, as well as the cost of storing those tapes in a vault offsite.

Some customers require incrementals or transaction logs to be subject to secondary operations, like replication or duplication. If such a requirement exists, it is important to narrow the requirement to the exact workloads or specific host clients that require incrementals and transaction logs to be replicated or duplicated. It is important not to paint this requirement with a broad brush because it seems expedient.

Another factor that many customers don't fully consider is the implication of an infinite retention. The idea of an infinite retention is not realistic.

Consider for a moment a large database containing patient records. That database should be protected by a variety of primary and secondary methods. Backup is a secondary method, whilst high availability and disaster recovery should be the primary protection methods. It is reasonable to require this database to maintain patient records infinitely, but the backups of that data shouldn't require an infinite retention. While most states and countries have requirements around record retention, it is important to understand those requirements and not just presume an infinite retention is required. It would be more reasonable for an LTR for monthly full backups to be up to 7 years. If the patient data is maintained in the database indefinitely, then that data would be in every single full backup taken. Why would a customer want to restore a database from greater than 30 days if the data in the source is never archived or purged? On the outside chance that the primary copy

of the database is corrupted, the business owner would want the most recent, consistent copy restored.

Keep this in mind when retention levels are set for n+1 copies of specific workloads.

Timing

Determining the timing of when backups must run based upon internal customer RPO, RTO, and SLAs is extremely important because it is a significant variable that drives when, and how fast backups and secondary operations like replication and duplication must be completed. When MSDP pools are unable to keep up with backup and secondary operations, that can result in missed backup windows and SLP backlog.

Consider a scenario when backups are kept for only 7 days, but the N+1 copy taken during secondary operations via an SLP is kept for 35 days. If backups don't meet their backup windows because of workload imbalance and/or solution under-sizing, then that performance degradation can impact the performance of secondary operations. In extreme situations, that can result in data that is being replicated and/or duplicated that is already considered past the retention period. Clearly, it is important to avoid that type of scenario, which is why sizing for compute, as well as capacity, is paramount.

Leveraging requirements and best practices

Design plans include many new features.

When the data gathering phase is complete, the next steps involve leveraging that data to calculate the capacity, I/O and compute requirements to determine three key numbers. Those numbers are BETB, IOPS, and compute (memory and CPU) resources. Veritas recommends that customers engage the Veritas Presales Team to assist with these calculations to determine the sizing of the solution. Then, it is important to consider some best practices around sizing and performance, as well as ensuring that the solution has some flexibility and headroom.

Best practice guidelines

Due to the nature of MSDP, the memory requirements are driven by the cache, spoold and spad. The guideline is 1GB of memory to 1TB of MSDP. For a 500TB MSDP pool, the recommendation is a minimum of 500GB of memory. Also note that leveraging features like Accelerator can be memory intensive. The memory sizing is important.

For the workloads that have very high job numbers, it is recommended that smaller disk drives be leveraged to increase IOPS performance. Sometimes 4TB drives are a better fit than 8TB drives. Consider this suggestion as a factor along with the workload type, data characteristics, retention, and secondary operations.

Where MSDP storage servers are virtual, whether through VMware, Docker, or in the cloud, it is important not to share physical LUNs between instances. Significant performance issues have been observed in MSDP storage servers that are deployed in AWS, Azure, VMware, and Docker when the physical LUNs are shared between instances.

Often, customers mistakenly believe that setting a high number of data streams on an MSDP pool can increase performance of their backups. However, the goal is to set the number of streams that satisfy the workload needs without creating a bottleneck due to too many concurrent streams fighting for resources. For example, a single MSDP storage server with a 500TB pool protecting Oracle workloads exclusively at 60K jobs per day was configured with a maximum concurrent stream count of 275. Initially, this count was set to 200 and then gradually increased to 275.

One method of determining if the stream count is too low, is to measure how long a single job, during the busiest times of the day, waits in queue. If a lot of jobs are waiting in the queue for lengthy periods, then it is possible the stream count is too low.

That said, it is important to gather performance data like SAR from the storage server to see how compute and I/O resources are used. If those resources are heavily used at the current state of a specific stream count, and yet there are still large numbers of jobs waiting in the queue for a lengthy period of time, then additional MSDP storage servers may be required to meet a customer's window for backups and secondary operations.

When it comes to secondary operations, the goal should be to process all SLP backlog within the same 24 hours it was placed in queue. As an example, if there are 40K backup images per day that must be replicated and duplicated, the goal is to process those images consistently within a 24-hour period to prevent a significant SLP backlog.

Customers often make the mistake of oversubscribing their Maximum Concurrent Jobs within their storage units (STUs). This mistake adds up to be a number larger than the Max Concurrent Streams on the MSDP pool. This approach is not a correct way to leverage STUs. Additionally, customers may incorrectly create multiple STUs that reference the same MSDP storage server with stream counts that individually aren't higher than the Max Concurrent Streams on the MSDP pool, but add up to a higher number when all STUs that reference that storage server are combined. This approach is also an improper use of STUs.

All actively concurrent STUs that reference a single, specific MSDP storage server must have Maximum Concurrent Jobs set in total to be less than or equal to the Maximum Concurrent Streams on the MSDP pool. STUs are used to throttle workloads that reference a single storage resource. For example, if Maximum Concurrent Streams for an MSDP pool is set to 200 and two storage units have

Maximum Concurrent Jobs each set to 150, the maximum number of jobs that can be processed at any given time is still 200, even though the sum of the two STUs is 300. This type of configuration isn't recommended. Furthermore, it is important to understand why more than one STU should be created to reference the same MSDP pool. A clean, concise NetBackup configuration is easier to manage and highly recommended. It is rare that a client must have more than one STU referencing the same MSDP storage server and associated pool.

Another thing to consider is that SLPs do need one or more streams to process secondary operations. Duplications and replications may not always have the luxury to be written during a window of time when no backups are running. Therefore, it is recommended that the sum of the Maximum Concurrent Jobs on all STUs referencing a specific MSDP storage server be 7-10% less than the Maximum Concurrent Streams on the MSDP pool to accommodate secondary operations while backups jobs are running. An example is where the Maximum Concurrent Streams on the MSDP pool is set to 275 while the sum of all Maximum Concurrent Jobs set on the STUs that reference that MSDP storage server is 250. This example allows up to 25 streams to be used for other activities like restores, replications, and duplications during which backups jobs are also running.

Pool Sizes

Although it is tempting to minimize the number of MSDP storage servers and size pools to the max 960TB, there are some performance implications that are worth considering. It has been observed that heavy mixed workloads sent to a single, 960TB MSDP pool don't perform as well as constructing two MSDP pools at 480TB and grouping the workloads to back up to a consistent MSDP pool. For example, consider two workload types, namely VMware and Oracle which happen to both be very large. Sending both workloads to a single large pool, especially considering that VMware and Oracle are resource-intensive, and both generate high job counts, can affect performance. In this scenario, creating a 480TB MSDP pool as the target for VMware workloads and a 480TB MSDP pool Oracle workloads can often deliver better performance.

Some customers incorrectly believe that alternating MSDP pools as the target or the same data is a good idea. It isn't. In fact, this approach decreases deduplication efficacy. Veritas does not recommend that a client send the same client data to two different pools. Also, Veritas does not recommended that a client send the same workloads to two different pools. This action negatively affects solution performance and capacity.

The only exceptions would be in the case that the target MSDP pool isn't available due to maintenance, and the backup jobs can't wait until it is available, or perhaps the MSDP pool is tight on space and juggling workloads temporarily is necessary while additional storage resources are added.

Fingerprint media servers

Many customers believe that minimizing the number of MSDP pools while maximizing the number of fingerprint media servers (FPMS) can increase performance significantly. In the past, there has been some evidence that FPMS might be effective at offloading some of the compute activity from the storage server would increase performance. While there are some scenarios where it might still be helpful, those scenarios are less frequent. In fact, often the opposite is true. There has been repeated evidence that large numbers of FPMSs leveraging a small number of storage servers can be a waste of resources, increase complexity, and affect performance negatively by overwhelming the storage server. We have consistently seen that the use of more storage servers with MSDP pools in the range of 500TB tend to perform better than a handful of FPMSs directing workloads to a single MSDP storage server. Therefore, it is recommended that the use of FPMS be deliberate and conservative, if they are indeed required.

Flexibility

The larger the pool, the larger the MSDP cache. The larger the pool, the longer it takes to run an MSDP check when the need arises. The fewer number of pools, the more the effect of taking a single pool offline for maintenance can have on the overall capability of the solution. Therefore, considering more pools of a smaller size instead of a minimum number of pools at a larger size can provide flexibility in your solution design and increase performance.

For virtual platforms such as Flex, there is value to creating MSDP pools and associated storage server instances that act as a target for a specific workload type. With multiple MSDP pools that do not share physical LUNs, the end result produces less I/O contention while minimizing the physical footprint.

Headroom

Customer who runs their environments very close to full capacity tend to put themselves in a difficult position when a single MSDP pool becomes unavailable for any reason. When designing a solution that involves defining the size and number of MSDP pools, it is important to minimize SPOF, whether due to capacity, maintenance, or component failure. Furthermore, in cases where there are a lot of secondary activities like duplications or replications, ensuring there is some additional capacity headroom is important, as certain types of maintenance activity might lead to a short-term SLP backlog. A guideline of 25% headroom in each MSDP pool is recommended for these purposes, whether SLP backlog or temporarily juggling workloads due to the aforementioned.

Cloud tier sizing and performance

Sizing and performance of data in the cloud is based on customer need and will vary from customer to customer, which makes providing exact sizing and performance information difficult.

To get data to the cloud, customers can use a simple internet connection if the data to be transmitted is less than the amount of available bandwidth.

Cloud providers may provide a direct connection service. Customers can get a dedicated link to their cloud provider with performance similar to LAN bandwidth inside a data center. They can compress data at the data center before being sent across the network to the cloud provider or use MSDP Cloud (MSDP-C) to optimize the data being sent before it is stored in the cloud. They can also throttle bandwidth, if desired, to prevent over-saturation of the network pipe.

Cloud instance model

Public cloud providers use a regional model in which it has configured various regions across the globe. Each region can have zones that are similar to data centers within the region that communicate with each other over high-bandwidth connections. This setup is similar to a customer having multiple physical data centers in a geographical region that are close enough for low-latency connectivity, yet far enough apart to not be affected by the same natural or artificial disaster.

Data within the region will typically stay within the region, but customers have the option to select a geographically dispersed region that is available for regional DR. Data can be replicated between zones to provide high availability within the cloud for the customer's data. The loss of a single zone does not affect the operations of the others. Customers typically choose to operate within the region that is closest to provide optimized bandwidth for moving data in and out of the Cloud and have the option to select a geographically dispersed region to provide regional DR.

Cloud storage options

One of the many benefits of the cloud storage model is the ability to quickly add storage to environments. Customers don't pay for the storage until it is provisioned. This model is much different from a traditional data center where racks of disks may sit idle until needed, thus increasing total cost of ownership (TCO). If the disk is spinning and generating heat, additional cooling and power could be needed to keep the disk spinning even if it is not currently in use. Although next-generation SSD arrays require less cooling and power, idle disks still increase TCO.

Once data is in the cloud, cloud providers uses various types of storage including block and object storage. Other options include Network-Attached Storage (NAS) services. Sizing of the environment is based on the needs of the customer and the workloads placed in the cloud. Pricing is based on the type of storage chosen and

is typically priced per GB. For example, standard object storage typically runs approximately $0.023/GB per month, whereas archive storage currently runs about $0.004/GB per month and even less for archive storage tiers. Cost also depends on the region where the data is stored. (See the cloud provider's price list for current costs.) Archive tiers are typically used as long-term archive storage targets whereby data is moved there automatically using a variety of methods, and a restore from storage could take hours to access versus seconds for a restore from object storage.

Environment description and assumptions for sizing

The following sample sizing guidelines are based on the assumptions listed and were created using the standard NetBackup Appliance Calculator to determine the storage required for each type of workload. This guideline applies to back up in the cloud workloads only.

The following assumptions were used to size this environment:

- Data assumptions:
  - Data split – 80% FS / 20% DB [no hypervisor level in the cloud]
  - Daily retention 2 weeks / weekly – 4 weeks / monthly 3 months
  - Daily change rate 2%, and year-over-year (YoY) growth 10% [sizing for 1 year only]
- Instance Type workload descriptions:
  - Small – FETB <=100 TB <= 100 concurrent jobs
  - Medium – FETB <=500 TB <= 500 concurrent jobs
  - Large – FETB <=1,000 TB <= 1,000 concurrent jobs
  - Extra-Large – FETB = 1 PB >1,000 concurrent jobs. For mixed object and block storage, total maximum capacity is 1.2PB

For more information, see *About the Media Server Deduplication (MSDP) node cloud tier* in the NetBackup Deduplication Guide.

NetBackup cloud instance sizing

The architecture is based on a single NetBackup domain consisting of a NetBackup primary server and multiple MSDP media servers in the cloud.

Typically, backups are written directly to local MSDP block storage for an immediate copy, then opt-duped to a cloud tier to send deduplicated data to object storage. However, there is no requirement that backups must go to standard MSDP before sending to cloud. If the solution doesn't require MSDP data to be "local" on block storage, then MSDP-C enables backup data to be sent directly to a cloud tier.

With MSDP-C or MSDP in general, data is deduplicated in the plug-in (source side) layer before writing to the storage server. The storage server assembles 64MB containers in memory and writes them to object storage. If the number of streams x 128MB exceeds the memory limits, disk is used as cache to assemble the containers.

Requirements consist of the following:

- Data assumptions:
  - NetBackup primary server
    A single NetBackup primary server can be on any supported operating system.
  - NetBackup MSDP media server's block storage
    MSDP media servers receive the initial backups from clients and perform deduplication.
  - NetBackup MSDP media server's cloud tier
    MSDP can have one or more targets on the same storage server that takes the deduplicated backup images from the MSDP media server's block storage and stores them in object storage.
    The MSDP cloud tier is dedicated to performing NetBackup deduplication writes to object storage. It is a dedicated high-end Red Hat server that meets the minimum requirements for the MSDP cloud tier. It takes the deduplicated backup images from the MSDP media servers and stores them in object storage.
  - Backup Workloads (Clients/Agents)
    These are the systems or applications that are being protected.

NetBackup primary server

The NetBackup primary server should be sized according to the standard Veritas guidelines depending on the load placed on the complete NetBackup domain. Plan accordingly for the initial needs of the environment. Cloud providers offer the added benefit of being able to scale up the systems as workloads grow. The solution can scale out by adding additional media server nodes.

Primary Server Memory and CPU Requirements

For details about the recommended memory and CPU processor requirements for the various environment sizes:

See Table 2-1 on page 26.

These estimates are based on the number of media servers and the number of jobs the primary server must support. You may need to increase the amount of RAM and number of processors based on other site-specific factors.

Primary Server Recommendations

Table 3-5 Primary server recommendations

Small	Medium	Large	Extra Large
32 GiB / 8 vCPU	64 GiB / 8 vCPU	64 GiB / 16 vCPU	128 GiB / 16 vCPU
Install 500 GB Catalog 5 GB	Install 500 GB Catalog 5 GB	Install 500 GB Catalog 10 GB	Install 500 GB Catalog 10 GB

NetBackup MSDP storage

NetBackup MSDP storage can reside on either a NetBackup Appliance, a Virtual Appliance or a build-your-own (BYO) virtual or physical host, including a cloud-based virtual instance. This section outlines MSDP in a public cloud Infrastructure-as-a-Service (IaaS) deployment.

The host computer's CPU and memory constrain how many jobs can run concurrently. The storage server requires sufficient capability for deduplication and for storage management. Processors for deduplication should have a high clock rate and high floating-point performance. Furthermore, high throughput per core is desirable. Each backup stream uses a separate core.

Table 3-6 Recommended specifications for MSDP media servers

Hardware component	MSDP media server specification
CPU	■ Veritas recommends at least a 2.2-GHz clock rate. A 64-bit processor is required. ■ At least 4 cores are required. Veritas recommends 8 cores. ■ For 64 TBs of storage, Intel x86-64 architecture requires 8 cores.
RAM	■ From 8 TB to 32 TB of storage, Veritas recommends 1 GB of dedicated RAM for 1 TB of block storage consumed. ■ However, beyond 32 TB storage, Veritas recommends more than 32 GB of RAM for better and enhanced performance. ■ MSDP-C uses a dynamic spooler cache based on previous and currently running backups and does not leverage the traditional persistent fingerprint pool. The size is set in the `MaxCloudCacheSize` parameter in `contentrouter.cfg`. The default setting is 20% ■ MSDP-C also will try and leverage memory as an upload/download cache before falling back on disk. This will be relative to the number of concurrent jobs and each job will use 128 MB of upload cache data. The default maximum value for `CloudUploadCacheSize` is 12 GB, which allows for roughly 90 concurrent jobs.

Table 3-6 Recommended specifications for MSDP media servers *(continued)*

Hardware component	MSDP media server specification
Storage	■ MSDP block storage will perform best with storage throughput of 250 MB/s or faster. Because many volumes/VMs have a 250 MB/s max, it's recommended to use a RAID0/1 stripe. ■ Start out with the expected needed storage based on deduplication rates. Storage can easily be expanded by adding additional volumes to MSDP. ■ MSDP-C does not use a dedicated cache volume. Rather, it will make non-persistent use of free storage on the MSDP server when needed. By default, MSDP-C does require at least 1 TB free space on the MSDP server per cloud tier, configurable in `contentrouter.cfg`.
Operating system	■ The operating system must be a supported 64-bit operating system. MSDP-C requires a RHEL/ CentOS 7.3 or later server. ■ See the operating system compatibility list at http://www.netbackup.com/compatibility.

Growing the media server

As the amount of data protected by a server increases, the load requirements on that host will increase. In that scenario, there is a simple solution. You can easily expand any virtual instance or add volumes. Refer to your cloud provider's documentation for information about resizing instances and adding storage to them.

Media Server Deduplication Pool (MSDP) recommendations

Running traditional MSDP in a cloud environment requires specific resources to be available such as a 10G network and volumes with provisioned IOPS. The recommendations below have been formulated using cloud provider kits that address MSDP pools of different sizes. These are just recommendations and specific customer environments may have different needs. Depending on the footprint, any of the environments below would work based on the sizes.

MSDP Considerations

An example MSDP storage pool size is up to 96 TB on Linux:

- Can be a direct backup target, use Fingerprinting Media Servers or a client-side deduplication target.
- MSDP will be storing all data on managed disks.

- The pool will be able to use optimized deduplication or Automated Image Replication (A.I.R.) to send images to any Veritas deduplication-compatible target, including MSDP-C.

Storage considerations

Although multiple deduplication storage servers can exist in a NetBackup domain, storage servers do not share their storage. Each storage server manages its own storage. Deduplication within a storage server is supported, including between different block and cloud Disk Pools. Optimized deduplication is supported between different storage servers in the same primary domain. Automated Image Replication (A.I.R.) is supported for replicating data and image metadata between primary domains. For a small 32 TB MSDP storage pool performing a single stream read or write operation, storage media 250 MB/sec is recommended for enterprise-level performance. Scaling the disk capacity to 250 TB recommends a performing 500 MB/sec transfer rate. Multiple volumes may be used to provision storage; however, each volume should be able to sustain 250 MB/sec of IO. Greater individual data stream capability or aggregate capability may be required to satisfy your objectives for simultaneous writing to and reading from disk. The suggested layout to use in a cloud infrastructure is a striped RAID0 or RAID1 configuration. More complex RAID configurations are not cost-efficient.

Table 3-7 shows recommended NetBackup media server sizing guidelines based on the size of the intended deduplication pool.

Table 3-7 Recommended media server sizing guidelines

Deduplication Pool	Storage	Cores	RAM in GB	Network in Gbps	IOPS
10 TB (Small)	1x160 SSD 1x16 TB SSD	8	64		
1–20 TB (Small)	1x80 SSD 1x16 TB SSD	8	32	10	
32 TB (Medium)	1x80 SSD 2x16 TB SSD	16	32	10	
	1x160 SSD 2x16 TB SSD IOPs – 12,000	36	64		12,000

Table 3-7 Recommended media server sizing guidelines *(continued)*

Deduplication Pool	Storage	Cores	RAM in GB	Network in Gbps	IOPS
32–64 TB (Large)	1x80 SSD 2-4x16 TB SSD	16	64	10	
	1x80 SSD 2-4x16 TB SSD	36	64	10	
	1x160 SSD 2x16 TB SSD IOPs – 12,000	40	160		12,000
32–64 TB (X-Large)	1x80 SSD 2-4x16 TB SSD	32	144	10	
	2x320 SSD 2-6x16 TB SSD IOPs -12,000	40	160	10	12,000

Table 3-8 Recommended initial NetBackup MSDP-C sizing guidelines

Role	Storage	CPUs	RAM
MSDP-C Small	250 GB SSD 1+TB	4	16 GB
MSDP-C Large	500 GB SSD 1+TB	8	32 GB

Accelerator performance considerations

NetBackup Accelerator uses platform, file system and application change detection logic to reduce the amount of data that needs to be read, transferred and written to backup storage this enable reduction of backup windows. Accelerator optimizes both full and incremental backups to reduce resource consumption (CPU, memory, storage IO, network IO) of backups on clients and media servers. As change detection and tracking capabilities differ based on the type of data being protected and the protection method, how NetBackup Accelerator works will vary by policy type and protection method.

NetBackup Accelerator requires backup storage that supports this capability. The storage needs to support the ability to create a new backup copy based on a prior backup (on that storage) by applying just the changes from the new accelerator

backup. Storage that is supported includes NetBackup Media Server Deduplication Pools (MSDP), NetBackup Cloud Connectors and OpenStorage partner storage solutions. Details of which storage solutions support accelerator are documented in the NetBackup Compatibility List for all Versions.

The first time a backup is run with Accelerator enabled a full backup will be run and the track log will be initialized. Accelerator optimization will not be seen until subsequent backups of that policy and client combinations to the same backup storage. When enabling Accelerator on a policy it is important to use the same storage unit in the policy and schedules to see the benefits of Accelerator optimization.

Accelerator for file-based backups

NetBackup Accelerator worsk with file and folder backup policy types Standard, Windows, and NDMP.

Track logs

NetBackup Accelerator uses files called track logs to enable tracking changes at a file and sub-file (segment) level on a policy/client basis. The track logs may be stored in different locations based on policy type and how a backup is performed. Individual track logs are created for each policy/client combination.

The track log size is relative to the size of the data (number and size of files) being backed up. So, enabling accelerator for backups of very large file servers or NAS may result in needing to manage the storage space for track logs.

The file system location where the track logs are stored needs sufficient space and performance characteristics for track log use. The track log is updated when an Accelerator backup is running. If the storage where the track log is located has high latency (read or write) or heavy I/O load, the performance of the backup may be negatively affected. Also, if multiple backups are running concurrently on the system (client or media server) where the track log is located that could be a source of I/O contention and should be monitored if there are performance concerns with Accelerator backups. Insufficient space for track logs may result in backup failures or the Accelerator optimization being disabled for the backup.

The following formula is used to compute track log size:

Track log size in bytes = (number of files * 128) + (((total disk space used) / 128KiB) * 20)

For example, if you have 179,418,964 files and 28,264,740,674 bytes in total disk space used, the result is:

(179418964*128) + (((28264740674/128)*20)) = 22965627392 + 4416365730.3125 = 27381993122.3125 bytes

More information is available about track log size:

See "Controlling disk space for Accelerator track logs" on page 80.

The following information also provides details on track log locations and information on how to estimate the space required for track logs:

- See *NetBackup logs for Accelerator* in the NetBackup Administrator's Guide, Voilume I
- How to calculate the NetBackup Accelerator track log size

Policy-specific information

Refer to the linked guides for detailed information about these policy types:

- Windows
  NetBackup Administrator's Guide, Volume I
- Standard
  NetBackup Administrator's Guide, Volume I
- NDMP
  NetBackup for NDMP Administrator's Guide

Controlling disk space for Accelerator track logs

At the start of an Accelerator-enabled backup, NetBackup can anticipate a disk-full situation due to track log processing. Accelerator track logs on a client can become a problem for clients with very little free space.

By default, NetBackup prevents an Accelerator backup if there is less than 5 GB or 5% of free space on the system.

NetBackup provides the two configuration settings to control the amount of free disk space on a host for an Accelerator backup to start:

- `ACCELERATOR_TRACKLOG_FREE_SPACE_MB`
- `ACCELERATOR_TRACKLOG_FREE_SPACE_PERCENT`

The default value for each setting is 5120 MB and 5%, respectively.

Note: If either of the settings is not specified in the configuration, the default value for that setting is used.

If you set the ACCELERATOR_TRACKLOG_FREE_SPACE_MB value to *X*, the job will not fail if there is more than *X* MB of free space and there is more than 5% disk space (the default for ACCELERATOR_TRACKLOG_FREE_SPACE_PERCENT).

Similarly, if you set the ACCELERATOR_TRACKLOG_FREE_SPACE_PERCENT value to *Y* will not fail a job if there is more than *Y*% disk space free and there is more than 5120 MB of disk space free (the default for ACCELERATOR_TRACKLOG_FREE_SPACE_MB).

To use only ACCELERATOR_TRACKLOG_FREE_SPACE_PERCENT and not ACCELERATOR_TRACKLOG_FREE_SPACE_MB, specify the ACCELERATOR_TRACKLOG_FREE_SPACE_PERCENT value and set the ACCELERATOR_TRACKLOG_FREE_SPACE_MB value to 0.

To use only ACCELERATOR_TRACKLOG_FREE_SPACE_MB and not ACCELERATOR_TRACKLOG_FREE_SPACE_PERCENT, specify the ACCELERATOR_TRACKLOG_FREE_SPACE_MB value and set the ACCELERATOR_TRACKLOG_FREE_SPACE_PERCENT value to 0.

Accelerator for virtual machine backups

In the case of virtual machine backups (VMware, Hyper-V, or RHV), change tracking is usually done using special APIs provided by hypervisors or storage stack, hence track logs are not used or maintained for such backups. Instead of a track log, an extent file is maintained which is much smaller in size compared to a files/folder track log. The extent file is maintained along with the rest of the state files used in case of these backups. This extent file is used to identify where disk extends for a VMware backup are mapped into the previous image, this information is crucial to identify how to reconstruct a new image combining changed data and unchanged data from the previous image.

The extent file has no file-level information, only extent (data offset and size) information.

Incremental Accelerator backups for Virtual Machines (VMware/Hyper-V) and Image Layout

VMware Accelerator supports Full and Incremental images. Full VMware backup generated with Accelerator has the same layout as Non-accelerator backups. Incremental backup image is different, it reads only changed blocks and only those are sent from the backup host to the media server, but a full image is synthesized on the server. This allows such incremental images to be used for operations such

as Instant Recovery / Instant Access. It also simplifies DR. On the flip side, as the size of the resultant image comparable to a full, which can have an adverse impact on tape-out performance. Replication performance for such images can also be impacted due to increased size, but MSDP and many other OST vendors take advantage of optimized duplication and impact is minimal in a homogeneous environment. There is no impact on the NetBackup catalog, the incremental image will catalog only those files that have changed. In other words, the catalog is the only key difference between Full and Incremental VMware backup images.

Forced rescan schedules

For virtualization technologies, such as VMware and Hyper-V, forced rescan schedules work differently than for files and folder backup. With files and folders, NetBackup maintains a track log to manage the change detection. However, with virtualization workloads, it is usually the virtualization vendor technology that provides the change detection.

Forced rescan mode in files and folder verifies information on the client file system with the information stored in the track log. For VMware backups, there is no track log, so there is no verification as such. Instead, forced rescan with virtualization workloads requests all blocks, regardless of any detected changes. This method ensures that all blocks are protected by that backup, which results in a full backup. The duration of the forced rescan backup takes longer and is similar to a non-Accelerator full backup.

For stable customer environments with proven technology stack, forced rescan schedules are not required for virtualization workloads. They can be used for troubleshooting and problem-solving purposes. For additional information on forced rescan for VMware, refer to the NetBackup fo VMware Adminstrator's Guide

Reporting the amount of Accelerator data transferred over the network

For Accelerator backup reporting, several NetBackup commands can report the amount of data that is transferred over the network for each Accelerator backup. The amount of transferred data is often much less than the size of the Accelerator backup image.

For backup reporting, it may be important to distinguish between the backup image size and the amount of data that was transferred over the network. You can configure the output of `bpimagelist`, `bpdbjobs`, and `bpclimagelist` to show the amount of Accelerator backup data that was transferred over the network instead of the backup image size.

Use `bpsetconfig` to set following configuration with an appropriate value:

```
REPLACE_IMAGE_SIZE_WITH_DATA_TRANSFERRED = <value>
```

Possible values:

- REPLACE_IMAGE_SIZE_FOR_ACCL_INC_VMWARE
- REPLACE_IMAGE_SIZE_FOR_ACCL_ALL_VMWARE
- REPLACE_IMAGE_SIZE_FOR_ACCL_INC_ALL
- REPLACE_IMAGE_SIZE_FOR_ACCL_ALL_ALL
- REPLACE_IMAGE_SIZE_FOR_ACCL_INC_HYPERV
- REPLACE_IMAGE_SIZE_FOR_ACCL_ALL_HYPERV
- REPLACE_IMAGE_SIZE_FOR_ACCL_INC_VIRTUAL
- REPLACE_IMAGE_SIZE_FOR_ACCL_ALL_VIRTUAL

Refer to the NetBackup for VMware Administrator's Guide for additional details and recommended procedures to set these values.

Accelerator backups and the NetBackup catalog

Use of Accelerator does not affect the size of the NetBackup catalog. A full backup with Accelerator generates the same catalog size as a full backup of the same data without Accelerator. The same is true of incremental backups: use of Accelerator does not require more catalog space than the same backup without Accelerator.

A potential catalog effect does exist, depending on how often you use Accelerator with full backups. A full backup with Accelerator completes faster than a normal full. It may therefore be tempting to replace your incremental backups with Accelerator full backups. Note: Since a full backup requires more catalog space than an incremental, replacing incrementals with fulls increases the catalog size. When changing your incrementals to fulls, you must weigh the advantage of Accelerator fulls against the greater catalog space that fulls require compared to incrementals.

Chapter 4

Media configuration guidelines

This chapter includes the following topics:

About dedicated versus shared backup environments

Your backup environment can be dedicated or shared.

Note the following:

- Dedicated SANs are secure but expensive.
- Shared environments cost less, but require more work to make them secure.
- A SAN installation with a database may require the performance of a RAID 1 array. An installation backing up a file structure may satisfy its needs with RAID 5 or NAS.

Suggestions for NetBackup media pools

The following are good practices for media pools (formerly known as volume pools):

- Configure a scratch pool for management of scratch tapes. If a scratch pool exists, EMM can move volumes from that pool to other pools that do not have volumes available.
- Use the available_media script in the goodies directory. You can put the available_media report into a script. The script redirects the report output to a file and emails the file to the administrators daily or weekly. The script helps track the tapes that are full, frozen, suspended, and so on. By means of a script, you can also filter the output of the available_media report to generate custom reports.
  To monitor media, you can also use the Veritas OpsCenter or Veritas APTARE. For instance, either one can issue an alert based on the number of media available or the percent of media that are frozen or suspended.
- Use the none pool for cleaning tapes.
- Do not create more pools than you need. In most cases, you need only 6 to 8 pools. The pools include a global scratch pool, catalog backup pool, and the default pools that are created by the installation. The existence of too many pools causes the library capacity to become fragmented across the pools. Consequently, the library becomes filled with many tapes that are partially full.

See "NetBackup job delays" on page 28.

See "Selection of storage units: performance considerations" on page 31.

Disk versus tape: performance considerations

Disk is now a common backup medium. Backup data on disk generally provides faster restores.

Tuning disk-based storage for performance is similar to tuning tape-based storage. The optimal buffer settings for a site can vary according to its configuration. It takes thorough testing to determine these settings.

Disk-based storage can be useful if you have a lot of incremental backups and the percentage of data change is small. If the volume of data in incremental copies is insufficient to ensure efficient writing to tape, consider disk storage. After writing the data to disk, you can use staging or storage lifecycle policies to copy batches of the images to tape. This arrangement can produce faster backups and prevent wear and tear on your tape drives.

Consider the following factors when backing up a data set to disk or tape:

- Short or long retention period
  Disk is well suited for short retention periods; tape is better suited for longer retention periods.
- Intermediate (staging) or long-term storage
  Disk is suited for staging; tape for long-term storage.
- Incremental or full backup
  Disk is better suited to low volume incremental backups.
- Synthetic backups
  Synthetic full backups are faster when incremental backups are stored on disk.
- Data recovery time
  Restore from disk is usually faster than from tape.
- Multistreamed restore
  Must a restore of the data be multistreamed from tape? If so, do not stage the multistreamed backup to disk before writing it to tape.
- Speed of the backups
  If client backups are too slow to keep the tape in motion, send the backups to disk. Later, staging or lifecycle policies can move the backup images to tape.
- Size of the backups
  If the backups are too small to keep the tape in motion, send the backups to disk. Small backups may include incrementals and frequent backups of small database log files. Staging or lifecycle policies can later move the backup images to tape.

The following are some benefits of backing up to disk rather than tape:

- No need to multiplex
  Backups to disk do not need to be multiplexed. Multiplexing is important with tape because it creates a steady flow of data which keeps the tape in motion efficiently (tape streaming). However, multiplexing to tape slows down a subsequent restore.
  More information is available on tape streaming.
  See "NetBackup storage device performance in the data transfer path" on page 210.
- Faster access to data
  Most tape drives have a "time to data" of close to two minutes. Time is required to move the tape from its slot, load it into the drive, and seek to an appropriate place on the tape. Disk has an effective time to data of 0 seconds. Restoring a large file system from 30 different tapes can add almost two hours to the restore: a two-minute delay per tape for load and seek, and a possible two-minute delay per tape for rewind and unload.

- Fewer full backups
  With tape-based systems, full backups must be done regularly because of the "time to data" issue. If full backups are not done regularly, a restore may require too many tapes from incremental backups. As a result, the time to restore increases, as does the chance that a single tape may cause the restore to fail.

NetBackup media not available

Some backup failures can occur because there is no media available. In that case, execute the following script and run the NetBackup Media List report to check the status of media:

Linux/UNIX

```
/usr/openv/netbackup/bin/goodies/available_media
```

Windows

```
install_path\NetBackup\bin\goodies\available_media
```

The NetBackup Media List report may show that some media is frozen and therefore cannot be used for backups.

I/O errors that recur can cause NetBackup to freeze media. The NetBackup Troubleshooting Guide describes how to deal with this issue. For example, see under NetBackup error code 96.

You can also configure the NetBackup error threshold value.

See “Adjusting the media_error_threshold” on page 88.

About the threshold for media errors

Each time a read, write, or position error occurs, NetBackup records the time, media ID, type of error, and drive index in the EMM database. Then NetBackup scans to see whether that media has had "*m*" of the same errors within the past "*n*" hours. The variable "*m*" is a tunable parameter known as the `media_error_threshold`. The default value of `media_error_threshold` is 2 errors. The variable "*n*" is the `time_window` parameter (default 12 hours).

If a tape volume has more than `media_error_threshold` errors, NetBackup takes the appropriate action.

Table 4-1 If number of tape volume errors exceeds media_error_threshold

Situation	NetBackup action
If the volume has not been previously assigned for backups	NetBackup does the following: ■ Sets the volume status to FROZEN ■ Selects a different volume ■ Logs an error
If the volume is in the NetBackup media catalog and was previously selected for backups	NetBackup does the following: ■ Sets the volume to SUSPENDED ■ Abandons the current backup ■ Logs an error

Adjusting the media_error_threshold

You can adjust the NetBackup media error threshold as follows.

To adjust the NetBackup media error thresholds

- Use the `nbemmcmd` command on the media server:

  Linux/UNIX

  ```
  /usr/openv/netbackup/bin/admincmd/nbemmcmd -changesetting
  -time_window unsigned integer -machinename string
  -media_error_threshold unsigned integer -drive_error_threshold
  unsigned integer
  ```

  Windows

  ```
  install_path\NetBackup\bin\admincmd\nbemmcmd.exe -changesetting
  -time_window unsigned integer -machinename string
  -media_error_threshold unsigned integer -drive_error_threshold
  unsigned integer
  ```

  For example, if the `-drive_error_threshold` is set to the default value of 2, the drive is downed after 3 errors in 12 hours. If the `-drive_error_threshold` is set to 6, it takes 7 errors in the same 12-hour period before the drive is downed.

  NetBackup freezes a tape volume or downs a drive for which these values are exceeded. For more detail on the `nbemmcmd` command, refer to the man page or to the *NetBackup Commands Reference Guide*.

About tape I/O error handling

Note: This topic has nothing to do with the number of times NetBackup retries a backup or restore that fails. That situation is controlled by the global configuration parameter `Backup Tries` for backups and the `bp.conf` entry `RESTORE_RETRIES` for restores.

The algorithm that is described here determines whether I/O errors on tape should cause media to be frozen or drives to be downed.

When a read/write/position error occurs on tape, the error that is returned by the operating system does not identify whether the tape or drive caused the error. To prevent the failure of all backups in a given time frame, `bptm` tries to identify a bad tape volume or drive based on past history.

To do so, `bptm` uses the following logic:

- Each time an I/O error occurs on a read/write/position, `bptm` logs the error in the following file.
  Linux/UNIX

  `/usr/openv/netbackup/db/media/errors`

  Windows

  `install_path``\NetBackup\db\media\errors`

  The error message includes the time of the error, media ID, drive index, and type of error. The following examples illustrate the entries in this file:

  ```
  07/21/96 04:15:17 A00167 4 WRITE_ERROR
  07/26/96 12:37:47 A00168 4 READ_ERROR
  ```

- Each time an entry is made, the past entries are scanned. The scan determines whether the same media ID or drive has had this type of error in the past "*n*" hours. "*n*" is known as the `time_window`. The default time window is 12 hours. During the history search for the `time_window` entries, EMM notes the past errors that match the media ID, the drive, or both. The purpose is to determine the cause of the error. For example: If a media ID gets write errors on more than one drive, the tape volume may be bad and NetBackup freezes the volume. If more than one media ID gets a particular error on the same drive, the drive goes to a "down" state. If only past errors are found on the same drive with the same media ID, EMM assumes that the volume is bad and freezes it.
- The freeze or down operation is not performed on the first error.

Note two other parameters: `media_error_threshold` and `drive_error_threshold`. For both of these parameters, the default is 2. For a freeze or down to happen, more than the threshold number of errors must occur. By default, at least three errors must occur in the time window for the same drive or media ID.

If either `media_error_threshold` or `drive_error_threshold` is 0, a freeze or down occurs the first time an I/O error occurs. `media_error_threshold` is looked at first, so if both values are 0, a freeze overrides a down. Veritas does not recommend that these values be set to 0.

A change to the default values is not recommended without good reason. One obvious change would be to put very large numbers in the threshold files. Large numbers in that file would disable the mechanism, such that to "freeze" a tape or "down" a drive should never occur.

Freezing and downing are primarily intended to benefit backups. If read errors occur on a restore, a freeze of media has little effect. NetBackup still accesses the tape to perform the restore. In the restore case, downing a bad drive may help.

For further tuning information on tape backup, see the following topics:

See "About the threshold for media errors" on page 87.

About NetBackup media manager tape drive selection

When NetBackup Enterprise Media Manager (EMM) determines which storage unit to use, it attempts to select a drive that matches the storage unit selection criteria. The criteria, for example, may be media server, robot number, robot type, or density.

Note the following:

- EMM prefers loaded drives over unloaded drives (a loaded drive removes the overhead of loading a media in a drive).
- If no loaded drives are available, EMM attempts to select the best usable drive that is suited for the job.
- In general, EMM prefers non-shared drives over shared drives, and it attempts to select the drive that was least recently used.

Chapter 5

How to identify performance bottlenecks

This chapter includes the following topics:

Introduction

Performance issues are different from other product defects and therefore the regular debugging process is not useful for debugging performance issues. When troubleshooting a product failure, the most common first step of debugging is to examine the default logs collected. If there is not enough information in the default logs, the next step is to elevate the logging level to collect more detail logs. For performance issue troubleshooting, elevating the log level is not recommended. This is because elevating the log level changes the system resource usage pattern and can substantially slow down the performance, masking the original problem and possibly creating a different performance issue.

This section discusses the correct approach for troubleshooting performance problems and covers the following sections:

- Proper mind set for performance issue RCA
- The 6 steps of performance issue RCA and resolution
- Flowchart of performance data analysis

Proper mind set for performance issue RCA

It is said that troubleshooting a performance issue is like looking for a needle in a haystack. The problem is vague and unstructured, moreover, it can be anywhere in the product and can be from both H/W components and software stack. Most non-performance engineers struggle with where to start the troubleshooting and many of them will dive into the area of their own expertise. For example, an FS expert will start at file system component, while a network engineer may start investigating the network layer. The mind set detailed in this section provides a structured approach to guide the resolution of an otherwise unstructured problem.

By following these guidelines, finding an entry point to start drilling down performance issue will become easier.

1. Block level understanding of both the hardware and the software components. Understanding the process flow to help narrow down the problem area.
2. Systematically drilling down the issue - top down and outside in, like peeling an onion. Always start by ensuring that the system has enough H/W resource bandwidth to handle the workload before jumping into the application tuning right away.
3. Tailor the tuning for each customer if necessary. Tuning that works for one customer may not work for the other, because differences in workload may create different tuning needs. So do not blindly apply a known tuning to other system unless the root cause is the same.
4. Be meticulous in data collection. Troubleshooting performance problems is an iterative process. As one bottleneck is resolved a new one may emerge. Therefore, automating data collection to ensure consistent data collection throughout the RCA process is critical for efficient problem resolution. In addition, avoid adding additional jobs or allowing unrelated jobs to run on the system while the data collection is in progress.
5. Remain relentless in RCA. Don’t attempt to tune the system until a root cause is identified. Without knowing the root cause, the tuning will be trial and error. It is time consuming and risky. Incorrect tuning can destabilize the system and can result in further performance degradation.
6. Keep laser focus on the four major resources – CPU, memory, IO, network. All performance issues manifest themselves in one or more of the 4 major H/W resources. By focusing on the usage patten of the four major resources, you can quickly identify an entry point to start the iterative RCA. Look for patterns that defy the common sense or the norm. For example, in general, higher throughput will consume more CPU cycles. If the throughput decreases, while CPU usage increases or remains the same, then your entry point should be the CPU. You may want to look for processes that consume more CPU. Another

example is when throughput has plateaued, but disk queue length increases. This is an indication of an I/O subsystem bottleneck and the entry point to RCA should be the I/O code path.

7. Performance numbers, both throughput and performance statistics, are relative. A number is meaningless until you compare with another number. For example, a disk queue length of 10 is meaningless until you compare with a similar workload which has a queue length of 5. That is why it is important to keep a set of performance data when system is running normally, and when a performance problem occurs, collect the same kind of data for comparison. Having a set of baseline numbers to compare with throughout the iterative process is key for successful problem resolution.
8. Identify changes in the Environment, such as newly implemented security requirements, changes in workloads applications, hardware of network infrastructure changes, and increases in size of data to the workloads.

The 6 steps of performance issue RCA and resolution

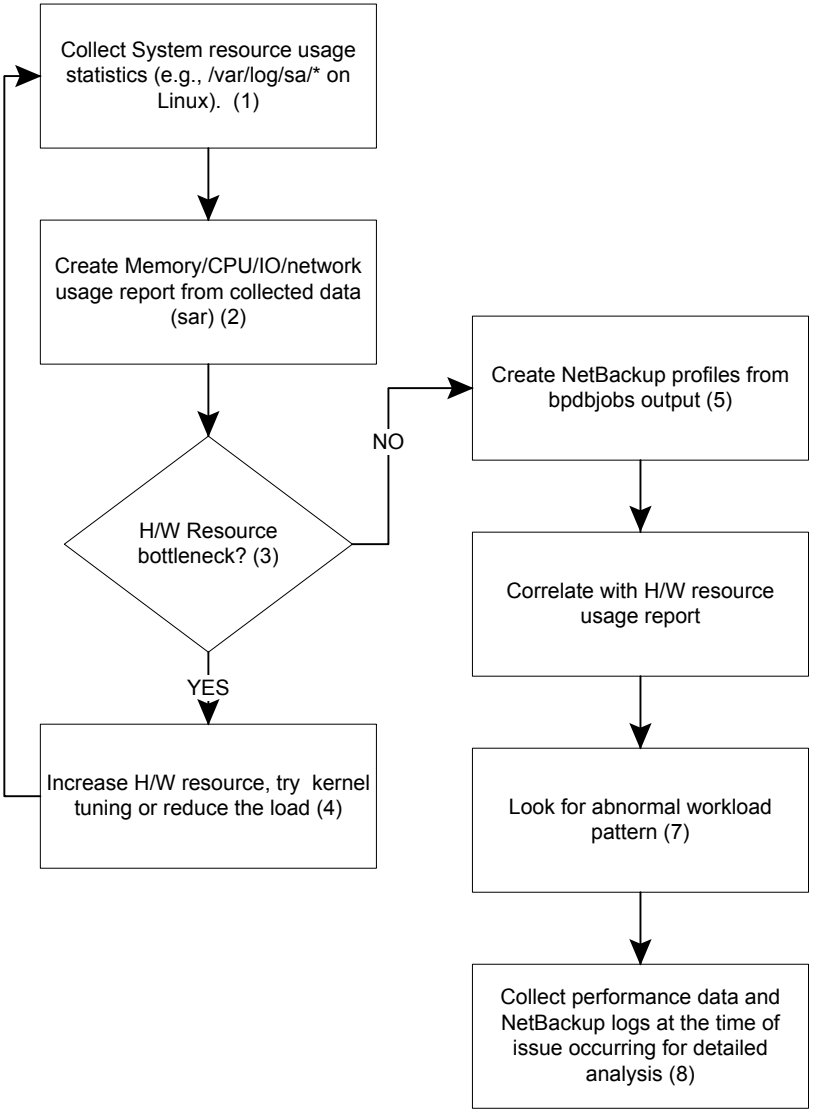
Use the following steps iteratively to drive the root cause analysis and resolution. Root cause analysis can be difficult but once the root cause is identified, the solution is usually trivial.

1. Obtain a clear problem description to ensure that the right problem is being solved.
2. Ensure the problem can be reproduced consistently. This step is important because if the problem can't be consistently reproduced, you never know whether the problem is resolved. Consistency also includes ensuring the result has minimum run-to-run variation. When run-to-run variation is more than 10 – 15%, it will be hard to determine if the improvement or degradation is due to tuning or just run-to-run variation.
3. Collect data for RCA. Data collection has to be meticulous. Avoid having other unrelated jobs or activities running while the data collection is in progress.
4. Perform root cause analysis. In particular, focus on the following two tasks. Use the following flow chart to guide the root cause analysis.
   - Analyze system resource usage
   - Create workload profile
5. Apply tuning based on the root cause identified in step 4.

6. If the problem persists, go back to step 2 again. Repeat until the problem is resolved.

Flowchart of performance data analysis

Figure 5-1 Flowchart of performance data analysis



NetBackup performance tuning can be grouped into three major categories: H/W resource availability, Kernel/OS tuning, and application, i.e., NetBackup tunings. RCA and tuning should be conducted in that order.

Jobs consume hardware (H/W) resources, in particular, the four major H/W resources: CPU, memory, I/O and network. As NetBackup servers age and the workload changes over time, hardware resource usage patterns will change also. Therefore, regularly monitoring/checking for the H/W resource usage pattern is important and should be the first step of any performance issue troubleshooting. This is because if the server is bottlenecked on any H/W resource, attempts to tune NetBackup to make it run faster will be futile, since the bottleneck will throttle the workload and prevent any further performance improvement. Therefore, the first step of RCA of a performance issue should always be checking for the H/W resource availability.

Boxes 1, 2, 3 and 4 in the above flowchart summarize the steps to detect and remove any H/W resource bottleneck. Some H/W resource shortages can be alleviated by kernel tunings. For example, on the Linux platform, frequent heavy swapping in the hundreds and thousands of KB per second is an indication of a memory bottleneck, thus, lowering the kernel parameter, swappiness, can reduce swapping without adding additional memory.

After all the H/W or kernel resource bottlenecks are eliminated, and the performance problem still exists, then most likely the root cause is in NetBackup. RCA and tuning of NetBackup should be the next step. Boxes 5, 6, 7 and 8 are the steps for finding the root cause of NetBackup bottlenecks.

How to create a workload profile

To profile the NetBackup workload, you can use the `bpdbjobs` command to collect the job detail report. The following is the command syntax to collect the job report:

```
/usr/openv/netbackup/bin/admincmd/bpdbjobs -report -all_columns >
/tmp/bpdbjobs.out
```

Note: To get a job summary, change the `-all_columns` option to `-most_columns`.

In the `bpdbjobs` report, many fields are included in each line of the job record. The following fields are useful to analyze the policy and job information.

Field	Description
`field2`	Job type (for example, backup, restore, duplicate, replication, and so on)
`field3`	State of the job (for example, queued, active, done, and so on)

Field	Description
`field8`	Media server that is used by the job
`field9`	Job started time
`field10`	Elapsed time for the job
`field11`	Job end time
`field22`	Policy type (for example Standard, Oracle, NDMP, VMware, and so on)
`field58`	Deduplication rate
`field61`	Deduplication ratio percent

More details of fields for the `bpdbjobs` command can be found in NetBackup Commands Reference Guide

Depending on the workload analysis target, the corresponding field can be used to analyze the jobs.

The following shows an example command to count the number of jobs for each type.

```
# awk -F, '{print $2}' /tmp/bpdbjobs.out | sort | uniq -c

7550 0 -> backup

9188 20 -> replication

8932 28 -> snapshot
```

Chapter 6

Best practices

This chapter includes the following topics:

- Best practices: Avoiding media server resource bottlenecks with Oracle VLDB backups
- Best practices: Avoiding media server resource bottlenecks with MSDPLB+ prefix policy
- Best practices: Cloud deployment considerations

Best practices: NetBackup SAN Client

The NetBackup SAN Client feature is designed for a computer that has the following characteristics:

- Contains critical data that requires high bandwidth for backups
- Is not a candidate for converting to a media server
- Has limited network bandwidth to use for backups.

A SAN Client performs a fast backup over a Fibre Channel SAN to a media server. Media servers that have been enabled for SAN Client backups are called Fibre Transport media servers.

Note: The FlashBackup option should not be used with SAN Clients. Restores of FlashBackup backups use the LAN path rather than the SAN path from media server to the client. The LAN may not be fast enough to handle a full volume (raw partition) restore of a FlashBackup backup.

More information is available in the NetBackup SAN Client and Fibre Transport Guide.

Best practices: NetBackup AdvancedDisk

With AdvancedDisk, NetBackup can fully use file systems native to the host operating system of the media server. NetBackup assumes full ownership of the file systems and also uses the storage server capabilities of the host operating system.

AdvancedDisk does not require any specialized hardware. AdvancedDisk disk types are managed as disk pools within NetBackup.

AdvancedDisk performance considerations

The entire data path between client and storage, including both hardware and software stacks, determines the overall performance of the backup and restore process. It is therefore essential that the performance of the disk storage is not

considered independently of entire data path and effect of infrastructure when seeking to resolve overall performance issues.

The AdvancedDisk storage implementation relates mounted files systems within the operating system of the media server to disk volumes within NetBackup disk pools. The normal rules and guidelines related to file system configuration on the host apply and advanced file systems such as VxFS, XFS, and ZFS can be used on media servers that support them. The AdvancedDisk storage implementation enforces the fact that an AdvancedDisk disk volume is a mounted file system on the storage server. However, NetBackup does not manage the mounting and dismounting of AdvancedDisk disk volumes. Instead, mounting and dismounting is managed by the administrator and the operating system of the host.

Recommendation: Configure file systems that are used as AdvancedDisk disk volumes in NetBackup in such a way that they are automatically mounted upon startup of the operating system of the hosting media server.

Exclusive use of disk volumes with AdvancedDisk

After a file system is imported as a disk volume into a NetBackup disk pool, NetBackup assumes that no third-party applications use this file system. However, NetBackup has no mechanism in place to enforce this assumption. Failure to ensure sole use or ownership of a disk volume by NetBackup may cause incorrect behavior of the capacity management components in NetBackup, resulting in premature image expiration.

For example, in a disk pool containing one disk volume with a high water mark set to 90% and a low water mark set to 70%, if 50% of the available space used by applications other than NetBackup only 40% will be available to NetBackup. When the high water mark is reached, 50% of all the backups held on the disk volume will be removed, rather than 20% if the disk volume is exclusively available to NetBackup.

Recommendation: You should not use disks that are used as AdvancedDisk volumes in NetBackup for any other purpose, including MSDP deduplication volumes.

Disk volumes with different characteristics

NetBackup allows multiple disk volumes to reside within a single disk pool. Load-balancing strategies are applied across all disk volumes within the disk pool during media and device selection (MDS). In other words, NetBackup assumes that all disk volumes within a disk pool are somewhat similar. This includes the areas of size and performance characteristics. While asymmetric configuration will not cause backups to fail, it is likely to result in unpredictable performance.

- Example 1:

Assume a disk pool with one 800-GB disk volume and one 80-GB disk volume. When NetBackup selects a disk volume to be used for a backup job, this selection is essentially driven by the size (free space) of the disk volume. Because of the size difference in this case, the larger disk volume will initially receive a larger share of the backup traffic. While this is quite correct from the point of view of available space, it will limit the system-wide performance as the bulk of the I/O traffic will be directed to the one disk.

- Example 2:
  Assume a disk pool with two disk volumes of the same size. However, one volume has a transfer rate of 100 MB/sec and the other has a transfer rate of 25 MB/sec. Once again, load balancing will select a disk volume based on available free space. As a result both disk volumes will see the same amount of traffic but backups and restores will run much faster to one disk volume than the other.

Recommendation: All disk volumes within a NetBackup disk pool should be of similar size and should have similar performance characteristics. When dealing with disk volumes that have significantly different characteristics they should be grouped into multiple disk pools, each containing disk volumes with similar characteristics, rather than all being placed in one pool.

Disk pools and volume managers with AdvancedDisk

Logical volume managers such Veritas Volume Manager (VxVM) allow abstractions to be created between the underlying disks/spindles and the volume on which a file system resides. Multiple small volumes can be created on a single disk and multiple disks can be combined to form a single large volume. Volume Managers can improve the resilience and data integrity for backups written to disk volumes used in disk pools by allowing mirroring and RAID configurations.

The AdvancedDisk storage implementation presents all mounted file systems as disk volumes to NetBackup. AdvancedDisk operates naturally with volume managers that work below the level of a mounted file system because it has no visibility of them and thus is not concerned with the underlying geometry of the storage.

Information about configuring disk pools is also available:

See “Best practices: Disk pool configuration - setting concurrent jobs and maximum I/O streams” on page 103.

Note: While the aforementioned guidance applies to both AdvancedDisk and MSDP pools, it is important to remember that AdvancedDisk does not use deduplication technology and it does not support the use of the NetBackup Accelerator feature.

Therefore, subsequent backup images written to the AdvancedDisk disk pool will not benefit from deduplication savings and increased performance provided by Accelerator and MSDP. Without the benefit of these features, backups usually require more capacity and typically take longer to complete. It is important to keep this in mind when setting Maximum I/O Streams and Maximum Concurrent Jobs. Usually, these values are set lower for AdvancedDisk pools than for MSDP pools.

While most environments will leverage MSDP and MSDP-C as storage targets for workload processing, there are some situations when AdvancedDisk is a better choice as a storage target:

- If the workload data characteristics do not deduplicate well, and there is no A.I.R. (replication) requirement.
- The workload is encrypted with 3rd party encryption and no A.I.R. (replication) requirement.
- Backup performance of database transaction logs or archive logs is increased when written to AdvancedDisk and there is no A.I.R. (replication) requirement.
- If A.I.R. is required, then the images backed up to the AdvancedDisk disk pool would need to be duplicated first to an MSDP or MSDP-WORM disk pool before being replicated or duplicated to another supported storage target.
- The use of the generic cloud connector as a way to duplicate images on AdvancedDisk to a supported cloud storage target can be rather slow. It is important to keep this in mind when considering the use of the generic cloud connector.

Network file system considerations

The AdvancedDisk storage implementation presents all mounted file systems as disk volumes to NetBackup. This includes network file systems, such as NFS and CIFS. Like different types of SCSI and SAN presented disk volumes, network file system disk volumes should be placed in dedicated disk pools based on size as well as performance. Disk pools should not contain a mixture of locally presented and network presented disk volumes.

Observe the following guidelines when using network file systems:

- Use manual mount points. Automatic mounting and dismounting can change mount points, which may cause disk resources to be unavailable.

- The server that exports the mount points must be configured to allow root access to the file systems from the storage servers.

If the NFS volumes are presented to more than one storage server, the following conditions must apply for the file systems of the disk volumes:

- Each media server must mount the file systems of all the disk volumes within a disk pool.
- The mount points must be valid.
- The mount points must be the same on each media server.

Some network file systems, such as NFS, implement a file system behavior that makes them unsafe for use in spanning situations.

Recommendation: Network file systems such as NFS should be configured to disable spanning.

State changes in AdvancedDisk

If the mount status of a configured disk volume changes, the state of the disk volume within NetBackup will follow after a delay of about a minute. You can check the current state of a disk volume with the command `nbdevquery -liststs -stype AdvancedDisk -U`, as this example from a Windows server shows:

```
Storage Server      : wstmas02
Storage Server Type : AdvancedDisk
Storage Type        : Formatted Disk, Direct Attached
State               : UP
Flag                : OpenStorage
Flag                : AdminUp
Flag                : InternalUp
Flag                : SpanImages
Flag                : LifeCycle
Flag                : CapacityMgmt
Flag                : FragmentImages
Flag                : Cpr
Flag                : RandomWrites
Flag                : FT-Transfer
```

The mount status is indicated by the `InternalUp` flag. If `InternalUp` is not shown, the disk volume is marked as `DOWN` in the output of the command `nbdevquery -listdp -U` as shown here:

```
Disk Pool Name      : DPtest
Disk Pool Id        : DPtest
Disk Type           : AdvancedDisk
Status              : DOWN
Flag                : Patchwork
Flag                : Visible
Flag                : OpenStorage
Flag                : SingleStorageServer
Flag                : AdminDown
Flag                : InternalUp
Flag                : SpanImages
Flag                : LifeCycle
Flag                : CapacityMgmt
Flag                : FragmentImages
Flag                : Cpr
Flag                : RandomWrites
Flag                : FT-Transfer
Raw Size (GB)       : 4.00
Usable Size (GB)    : 4.00
Num Volumes         : 1
High Watermark      : 98
Low Watermark       : 80
Comment             :
Storage Server      : wstmas02
```

The administrator can reset the state of the disk volume by issuing the command:

```
nbdevconfig -changestate -stype AdvancedDisk -dp <disk pool> -dv
<disk volume> -state RESET
```

Best practices: Disk pool configuration - setting concurrent jobs and maximum I/O streams

Setting Maximum Concurrent Jobs and Maximum I/O Streams are an important means of throttling workloads where both AdvancedDisk or an MSDP disk pool is the target.

- The correlation between Maximum Concurrent Jobs set on the storage unit, and Maximum I/O Streams set on the disk pool is significant. The Maximum I/O Streams setting on the disk pool, whether MSDP or AdvancedDisk, sets the maximum number of streams that can run concurrently on that disk pool,

regardless of the Maximum Concurrent Jobs setting on the storage unit, even if multiple storage units exist referencing the same disk pool.

- In the case that there are multiple storage units that reference the same disk pool, the total number of concurrent jobs that is configured to run is equal to the total of the Maximum Concurrent Jobs that is set across each storage unit referencing that specific disk pool. Whatever that total Maximum Concurrent Jobs number is, the maximum jobs that can actually be run is a number that is less than or equal to the Maximum I/O Streams set on the disk pool.
- For MSDP pools, it is recommended that the Maximum Concurrent Jobs sum should not be greater than 90% of the Maximum I/O Streams set on the disk pool, or by subtracting 10 from the total, whichever is greater.
- For example, a single MSDP pool with Maximum I/O Streams set to 200 should have a total Maximum Concurrent Jobs value not greater than 180. If more than one storage unit references that MSDP pool, the sum across all those storage units should be 180 or less. Meaning that in this example, a single storage unit referencing that disk pool would have Maximum Concurrent Jobs set to 180. If there were two storage units that referenced that same MSDP pool, then that total Maximum Concurrent Jobs could be 100 set on one and 80 set to the other, or any combination totaling no more than 180.
- In the case of AdvancedDisk disk pools, we recommend staying with the 90% recommendation, as that still reserves some streams for duplication.
- For example, in the case of an AdvancedDisk pool with Maximum I/O Streams set to 30, the Maximum Concurrent Jobs should be set to no more than 27.
- That said, to minimize complexity, we recommend that the use of multiple storage units referencing the same disk pool is limited, and ideally avoided.
- The 10% of the Maximum I/O Streams that is reserved is specifically used for secondary operations, like the use of A.I.R. with MSDP, or duplication activities which are supported with both AdvancedDisk and MSDP.
- One of the most common backup performance issues is caused by a mismatch between the Maximum I/O Streams and Maximum Concurrent Jobs. This configuration error often results in throughput bottlenecks that can prevent workloads from completing within their allotted backup windows.
- When the Maximum I/O Streams is set on the disk pool, the default is set to 2 per volume. It is recommended that you set this value to 10 I/O streams per independent LUN. For example, if you have 6 LUNs per disk pool, set the value to 10 and the total Maximum I/O streams would be 60 (6 x 10). However, a disk pool created for MSDP works differently. The entire disk pool is considered to be one volume regardless of the number of LUNs included. Therefore, for a 6 LUN MSDP disk pool, the value should be set to 60. .

- The recommended way to determine the ideal setting for Maximum I/O Streams is to start low (proportional to the number of volumes in the disk pool), monitor compute, I/O, and network bandwidth, and increase to the point that jobs are not waiting in queue for more than a minute before becoming active, and while ensuring that compute, I/O, and network bandwidth does not become depleted. Remember when the Maximum I/O Streams setting is changed on a disk pool, that Maximum Concurrent Jobs should be updated based upon the aforementioned guidance. Try to keep the aggregate media server hardware resource usage under 75%.
- It is not recommended that Maximum I/O Streams ever be set to unlimited.

More information is available in the following technical article:

NetBackup - Configuring the Max Jobs and Max IO Streams settings

Best practices: About disk staging and NetBackup performance

Duplicating to disk or tape from AdvancedDisk or MSDP is supported. However, when the source images reside on an MSDP pool, the rehydration process can lengthen the time required to complete the duplication process, specifically when the aforementioned images are duplicated to tape or to another storage target that does not support the use of deduplication technology.

That said, if the backup images reside on MSDP, then duplicating to another disk or S3 storage target which support direct duplication of deduplicated images without rehydration will typically be much faster. The aforementioned process is referred to as Optimized Duplication or Opt-Dup.

Best practices: Supported tape drive technologies for NetBackup

Recent tape drives offer noticeably higher capacity than the previous generation of tape drives that are targeted at the open-systems market. Administrators can take advantage of these higher-capacity and higher-performance tape drives. Consult the following documents for the latest information about supported tape drives and other devices:

Veritas NetBackup Compatibility List for all Versions

Veritas NetBackup for NDMP: NAS appliance information

Best practices: NetBackup tape drive cleaning

You can use the following tape drive cleaning methods in a NetBackup installation:

- Frequency-based cleaning
- TapeAlert (on-demand cleaning)
- Robotic cleaning

Refer to the *NetBackup Administrator's Guide, Volume II*, for details on how to use these methods. Following are brief summaries of each method.

Table 6-1 describes the three tape drive cleaning methods.

Table 6-1 Tape drive cleaning methods

Tape drive cleaning method	Description
Frequency-based cleaning	NetBackup performs frequency-based cleaning by tracking the number of hours a drive has been in use. When this time reaches a configurable parameter, NetBackup creates a job that mounts and exercises a cleaning tape. This practice cleans the drive in a preventive fashion. The advantage of this method is that typically no drives are unavailable awaiting cleaning. No limitation exists as to the platform type or robot type. On the downside, cleaning is done more often than necessary. Frequency-based cleaning adds system wear and takes time that can be used to write to the drive. This method is also hard to tune. When new tapes are used, drive cleaning is needed less frequently; the need for cleaning increases as the tape inventory ages. This method increases the amount of tuning administration that is needed and, consequently, the margin of error.

Table 6-1 Tape drive cleaning methods *(continued)*

Tape drive cleaning method	Description
TapeAlert (reactive cleaning, or on-demand cleaning)	TapeAlert (on-demand cleaning) allows reactive cleaning for most drive types. TapeAlert allows a tape drive to notify EMM when it needs to be cleaned. EMM then performs the cleaning. You must have a cleaning tape configured in at least one library slot to use this feature. TapeAlert is the recommended cleaning solution if it can be implemented. Certain drives at some firmware levels do not support this type of reactive cleaning. If reactive cleaning is not supported, frequency-based cleaning may be substituted. This solution is not vendor or platform specific. Veritas has not tested the specific firmware levels. The vendor should be able to confirm whether the TapeAlert feature is supported. See "How NetBackup TapeAlert works" on page 107. See "Disabling TapeAlert" on page 108.
Robotic cleaning	Robotic cleaning is not proactive, and is not subject to the limitations of the other drive cleaning methods. Unnecessary cleanings are eliminated, and frequency tuning is not an issue. The drive can spend more time moving data, rather than in maintenance operations. NetBackup EMM does not support library-based cleaning for most robots, because robotic library and operating systems vendors implement this type of cleaning in different ways.

How NetBackup TapeAlert works

To understand drive-cleaning TapeAlert, it is important to understand the TapeAlert interface to a drive. The TapeAlert interface to a tape drive is by means of the SCSI bus. The interface is based on a Log Sense page, which contains 64 alert flags. The conditions that cause a flag to be set and cleared are device-specific and device-vendor specific.

The configuration of the Log Sense page is by means of a Mode Select page. The Mode Sense/Select configuration of the TapeAlert interface is compatible with the SMART diagnostic standard for disk drives.

NetBackup reads the TapeAlert Log Sense page at the beginning and end of a write or read job. TapeAlert flags 20 to 25 are used for cleaning management, although some drive vendors' implementations vary. NetBackup uses TapeAlert

flag 20 (Clean Now) and TapeAlert flag 21 (Clean Periodic) to determine when to clean a drive.

When NetBackup selects a drive for a backup, `bptm` reviews the Log Sense page for status. If one of the clean flags is set, the drive is cleaned before the job starts. If a backup is in progress and a clean flag is set, the flag is not read until a tape is dismounted from the drive.

If a job spans media and, during the first tape, one of the clean flags is set, the following occurs: the cleaning light comes on and the drive is cleaned before the second piece of media is mounted in the drive.

The implication is that the present job concludes its ongoing write despite a TapeAlert Clean Now or Clean Periodic message. That is, the TapeAlert does not require the loss of what has been written to tape so far. This implication is true regardless of the number of NetBackup jobs that are involved in writing out the rest of the media.

If a large number of media become FROZEN as a result of having implemented TapeAlert, other media or tape drive issues are likely to exist.

Disabling TapeAlert

Warning: You should disable TapeAlert only during diagnostic testing. Disabling this feature can allow a tape device to write corrupt data during the backup thereby rendering the data useless for restore, import, and duplication.

Use the following procedure to disable TapeAlert.

To disable TapeAlert in NetBackup

- Create a touch file called `NO_TAPEALERT`, as follows:

  UNIX:

  ```
  /usr/openv/volmgr/database/NO_TAPEALERT
  ```

  Windows:

  ```
  install_path\volmgr\database\NO_TAPEALERT
  ```

Best practices: NetBackup data recovery methods

Recovering from data loss involves both planning and technology to support your Recovery Point Objectives (RPO) and Recovery Time Objectives (RTO) and time frames. You should document your methods and procedures and test them regularly to ensure that your installation can recover from a disaster.

Table 6-2 describes how you can use NetBackup and other tools to recover from various mishaps.

Table 6-2 Methods and procedures for data recovery

Operational risk	Recovery possible?	Recovery methods and procedures
File deleted before backup	No	None
File deleted after backup	Yes	Standard NetBackup restore procedures
Backup client failure	Yes	Data recovery using NetBackup
Media failure	Yes	Backup image duplication: create multiple backup copies
Media server failure	Yes	Recovery is usually automatic, by preconfiguring storage unit groups (for backup) and `FAILOVER_RESTORE_MEDIA_SERVER` (for restore from tape).
Primary server failure	Yes	Deploy the primary server in a cluster, for automatic failover
Loss of backup database	Yes	NetBackup database recovery
Complete site disaster	Yes	Vaulting and off-site media storage

Best practices: Suggestions for disaster recovery planning

You should have a well-documented and tested plan to recover from a logical error, an operator error, or a site disaster.

Information about disaster recovery is available in the following documents, which are found on the NetBackup Documentation Landing Page:

NetBackup Troubleshooting Guide (See *Chapter 4: Disaster recovery*)

NetBackup Administrator's Guide, Volumes I & II

NetBackup in Highly Available Environments Administrator's Guide

NetBackup Clustered Primary Server Administrator's Guide

NetBackup Cloud Administrator's Guide

For recovery planning, use the following preparatory measures:

- Always use a scheduled catalog backup.
  Refer to *Catalog Recovery from an Online Backup* in the NetBackup Troubleshooting Guide.
- Review the disaster recovery plan often.
  Review your site-specific recovery procedures and verify that they are accurate and up-to-date. Also, verify that the more complex systems, such as the NetBackup primary and media servers, have procedures for rebuilding the computers with the latest software.
  Performance depends on a chain of I/O events. The speed is only as good as the weakest link in that chain. For data recovery, you must consider I/O performance on the source, locality of the data, rehydration performance, and network performance.
  A number of NetBackup features optimize recovery, such as Instant Access and Instant Rollback for VMware.
  For MSDP-C the best recovery guidance is to have a reasonable amount of free space on the storage to use as a recovery cache.
- Perform test recoveries on a regular basis.
  Implement a plan to perform restores of various systems to alternate locations. This plan should include selecting random production backups and restoring the data to a non-production system. A checksum can then be performed on one or many of the restored files and compared to the actual production data. Be sure to include off-site storage as part of this testing. The end user or application administrator can also help determine the integrity of the restored data.
- Use and protect the NetBackup catalog.

  Do the following:

  - Back up the NetBackup catalog.
    The catalog contains information vital for NetBackup recovery. Its loss can result in hours or days of recovery time using manual processes.
  - Back up the catalog after each backup.
    If a catalog backup is used, an incremental catalog backup can be done after each backup session. Busy backup environments should also use a scheduled catalog backup, because their backup sessions end infrequently.
    In the event of a catastrophic failure, the recovery of images is slow if some images are not available. If a manual backup occurs shortly before the primary server or the drive that contains the backed-up files crashes, the manual backup must be imported to recover the most recent version of the files.

- Regularly confirm the integrity of the NetBackup catalog.
  Walk through the process of recovering a catalog from the device that contains the catalog backup on a regular basis, such as quarterly or after major operational or personnel changes. This essential part of NetBackup administration can save hours in the event of a catastrophe.

Best practices: NetBackup naming conventions

Use a consistent name convention on all NetBackup primary servers. Use lower case for all names. Case-related issues can occur when the installation comprises UNIX and Windows primary and media servers.

Table 6-3

Object to name	Guidelines for naming
Policy	One good naming convention for policies is *platform_datatype_server(s)*. Example 1: w2008_filesystems_trundle This policy name designates a policy for a single Windows 2008 server doing file system backups. Example 2: w2008_sql_servers This policy name designates a policy for backing up a set of Windows 2008 SQL servers. Several servers may be backed up by this policy. Servers that are candidates for being included in a single policy are those running the same operating system and with the same backup requirements. Grouping servers within a single policy reduces the number of policies and eases the management of NetBackup.
Schedule	Create a generic scheme for schedule names. One recommended set of schedule names is daily, weekly, and monthly. Another recommended set of names is incremental, cumulative, and full. This convention keeps the management of NetBackup at a minimum. It also helps with the implementation of Vault, if your site uses Vault.
Storage unit and storage group	A good naming convention for storage units is to name the storage unit after the media server and the type of data being backed up. Two examples: `mercury_filesystems and mercury_databases` Where "*mercury*" is the name of the media server, and "*file systems*" and "*databases*" identify the type of data being backed up.

Best practices: NetBackup duplication

Note the following about NetBackup image duplication:

- When duplicating an image, specify a volume pool that is different from the volume pool of the original image. (Use the **Setup Duplication Variables** dialog of the NetBackup Administration Console.)
- If multiple duplication jobs are active simultaneously (such as duplication jobs for Vault), specify a different storage unit or volume pool for each job. Using different destinations may prevent media swapping among the duplication jobs.
- NetBackup provides the `bpduplicate` command to run and script duplication jobs. Veritas, however, recommends using either storage lifecycle policies or the Vault option when you implement duplication as part of a backup strategy.

Best practices: NetBackup deduplication

Veritas recommends that you use the NetBackup stream handler where it is available for backup data streams. For data streams without the stream handler, NetBackup uses the default fixed-length deduplication (FLD) of 128 KB to segment the data streams for fingerprint calculation.

If the deduplication rates are poor with the default FLD, consider the following options:

- Try variable length deduplication (VLD). VLD can substantially improve the deduplication ratio for some database workload backups such as DB2, for which the stream handler is not available. A customer reported that the deduplication ratio improved from 0% to over 80% after switching to VLD.
  More information about how VLD works and the resource overhead is available:
  See "About NetBackup Media Server Deduplication (MSDP)" on page 54.
- Use AdvancedDisk. If VLD fails to increase the deduplication ratio, consider AdvancedDisk. Internal test results showed that backup to AdvancedDisk can outperform a 0% deduplication rate for backups by up to 25%.

For VLD, the default window size for scanning is 32KB - 128KB. We recommend that you use the default window size first. If you see improved deduplication ratios but want to fine-tune the window size to improve performance, you can use data from the job detail report from `bpdbjobs` command for analysis. The following is a sample `bpdbjobs` command syntax to extract detail DB2 job information:

```
bpdbjobs -report -all_columns -jobid db2jobid | tr "," "\n"
```

The following output shows that out of the 99.5% deduplication ratio, 96.3% comes from scanning the `104K-128K` segment size. By giving up 3.6% (1.2 + 1.3 + 1.1) of

the deduplication ratio, the scanning window size can be substantially reduced. This reduction can result in a substantial CPU saving because fingerprinting is a CPU-intensive operation.

```
VLD enabled\
SO Count=1417216\
32K~56K:1.2%\
56K~80K:1.3%\
80K~104K:1.1%\
104K~128K:96.3%) for (full qualified media server hostname ): scanned: 17811
CR sent: 931507 KB\
CR sent over FC: 0 KB\
dedup: 99.5%\
cache hits: 179 (0.0%)\
rebased: 367 (0.0%)\
where dedup space saving:98.9%\
compression space saving:0.5%
```

For more NetBackup deduplication best practices, see the *MSDP deployment best practices* section in the NetBackup Deduplication Guide.

Best practices: Universal shares

NetBackup's universal shares feature provides open-source data protection without the need for a backup agent or API. The feature was designed for use with Oracle, MS-SQL, and other relational database types so that a database analyst (DBA) can "dump" a database backup to a specific network share that is mounted on the database client via NFS or SMB. At that point, the "dump" can then be protected with a specialized NetBackup policy. This approach provides the DBA the flexibility to manage their own backups while also realizing the existing protection and optimization benefits of other NetBackup features like MSDP.

Space efficiency is achieved by storing this data directly into an existing NetBackup-based deduplication pool. Any data that is stored in a universal share is automatically placed in MSDP storage where it is deduplicated automatically. This data is then deduplicated against all other data that was previously ingested into the media server's MSDP pool. Because a typical MSDP storage server stores data across a broad scope of data types, the universal share offers significant deduplication efficiency.

After the data is protected and indexed within the NetBackup catalog, the data can be duplicated or replicated as part of any activity that is supported by a NetBackup storage lifecycle policy (SLP).

Benefits of universal shares

NetBackup's universal share feature provides flexibility, scalability, ease of recoverability to DBAs, all while providing the benefits of existing space-optimized data protection with MSDP. The feature provides support for millions of files per share, and scalability for multi-TB databases. These shares use the same deduplication pool as existing backups. Therefore, all data, whether from the shares or direct to the pool, realizes deduplication benefits. Universal share data is written to the same MSDP location that is used by all other backup activities. All data that is sent to an MSDP-based storage unit or to a universal share will coexist in the same MSDP pool. This optimization is transparent to DBAs.

Protection points facilitate data persistence, data retention, and indexing of the data within the NetBackup primary catalog. They provide single file search and recovery, as well as the ability to conduct secondary operations like replication and optimized duplication. Recovery is flexible. It does not affect existing universal shares. The data is referenced by share, not by media server. Even though the central MSDP pool is used for all shares on a media server, data that is placed in any given share is not visible or accessible by any other share.

Although the universal share feature has existed in previous versions, it required the use of a NetBackup Appliance. However, that support has expanded in NetBackup 9.1 to include build-your-own (BYO) media servers. The universal share feature is supported on an MSDP BYO storage server with NetBackup 9.1 and Red Hat Enterprise Linux 7.6 and later.

Furthermore, support for universal shares has been extended to the NetBackup Flex Appliance platform as of Flex version 2.0.1 while running NetBackup application instances at version 9.1 or later.

Configuring universal shares

To optimize the use of this feature, it is important to consider some key points directly related to configuration. How the shares are configured directly affects scalability and performance.

Scalability

As a guideline, it is recommended that no more than 50 shares be created per NetBackup media server or NetBackup Flex Instance. This recommendation is a guideline only and not a hard limit. That said, significant performance testing has revealed that performance can be affected when surpassing more than 50 concurrent shares. For clarity, the term "concurrent" in this context refers to active executing read and write operations. It was also observed that performance tends to peak at 25 concurrent shares.

To provide the most flexibility, leveraging the NetBackup Flex Appliance provides a way to create multiple MSDP instances, with the optimal 50 shares per MSDP instance.

As with all solution design, it is important to be mindful of the amount of compute and I/O resources available on the target hardware. Furthermore, all best practice recommendations around optimizing MSDP performance still apply here as the underlying technology on the storage target is MSDP. The recommendation of 1 GB of memory for 1 TB of MSDP storage still applies here as well.

When leveraging Flex Appliances with universal shares, the same principles to avoiding I/O bottlenecks apply. For example, avoid sharing LUNs across MSDP instances.

Best practices for Flex Appliances, traditional NetBackup Appliances, and BYO still apply as the universal share feature leverages the same underlying MSDP technology.

Universal share size is limited to 960 TB.

Host-to-share mapping

Each individual share can be used by multiple hosts. However, it is recommended that one share not be assigned to more than a few host clients, especially if each host client is frequently dumping data to the share. A share that is mapped to many host clients can experience performance bottlenecks that affect the success of universal share backups and secondary operations that are executed thereafter. For very busy environments, a 1:1 ratio of share to host client is optimal.

Universal share backups

Any data that is ingested into the universal share resides in the MSDP storage pool that is located on the appliance-based or BYO media server hosting the universal share. While any data ingested into the universal share is deduplicated and located in MSDP immediately, that data will not be referenced in the NetBackup catalog and no retention enforcement enabled before running a universal share backup. Without a universal share backup, the data that is placed in the universal share is not searchable and cannot be restored using standard NetBackup procedures. Before the backup, control of the data in the share is entirely managed by the host that is mounting the share. If the owner of the share deletes the share data or if the share itself is removed, the data that used to exist in the share is not recoverable by NetBackup. Therefore, the universal share protection point backup, a special backup type, was designed to facilitate management and restorability through traditional NetBackup methods.

For clarity, references to a universal share backup and a universal share protection point are the same in that they both refer to the special NetBackup policy type that

indexes the data in the share and sets the retention enforcement, making it available for other activities like secondary operations.

A single NetBackup policy can be configured to protect every universal share within a NetBackup domain or multiple NetBackup policies can be configured to protect each individual universal share. When a protection point is executed, no data movement occurs. Furthermore, the performance of this special backup is not based on the size of the file data. It is more closely correlated with the number of files in the specific universal share. As part of the special backup activity, each file in the share is indexed within the NetBackup catalog, and retention enforcement is set.

The timing of a universal share protection point backup is important for two important reasons:

1. It is important to ensure that the database dump is completed before initiating the protection point backup. Performance suffers if the backup is run while the database dump is still in progress. It can also affect how complete the backup is.
2. It is important for NetBackup administrators to meet with the DBAs to understand the workload size by host client, dump frequency, and time that is required to complete the dump. This information helps determine the optimal quiet period to schedule the backup of each share, as well as any subsequent secondary operations like replication and optimized duplication.

Running a universal share protection point backup during the quiet period when no dumps are occurring on the share helps to ensure that the complete dump is captured, as well as avoiding I/O contention between extensive read and write activities.

In reference to the recommendation of the optimal 1:1 ratio of host client to share mapping and scheduling the backup and any secondary operations during a quiet period, the 1:1 ratio helps prevent a scenario where there are too many host clients hitting a specific share, thus making it difficult to find a quiet period, as well as creating inevitable I/O contention.

The results of extensive testing where each NetBackup protection point policy backs up a small number of shares, for example, ~10 shares, and where each host client is mapped to one share, were favorable and allowed time for secondary operations.

It is also important to note that the NetBackup Accelerator feature does not apply here, nor is it supported.

Secondary Operations

Any functionality that is available with storage lifecycle policies (SLP) can be applied to data managed by a universal share protection point backup. This functionality includes transitioning data to tape, cloud, optimized duplication (opt-dup) to other

media servers, and replication to other NetBackup domains via Auto Image Replication (A.I.R.).

The maximum 50 concurrent universal share guideline includes read and write activities, including secondary operations.

To optimize performance of secondary operations, schedule these activities when no other read and write activities to the same share are occurring. For example, after the dump and the backup are completed.

Data characteristics

As previously highlighted, the data characteristics that affect deduplication efficacy also apply here as the underlying technology is MSDP. If a DBA chooses to use third-party encryption with their database dumps, the deduplication rate will be affected negatively. Data leveraging third-party encryption doesn't deduplicate well. Furthermore, certain types of database dump compression can also negatively affect deduplication efficacy. In both cases, decreased deduplication efficacy negatively affects space optimization, and it will also affect the speed requirements and the storage requirements of secondary operations.

It is also important to note that data characteristics where the dumps universal share consist of millions of tiny files will also be affected due to the overhead in read and write activities.

For all the aforementioned data characteristics, it is important run some real performance benchmark tests to measure speed and deduplication efficacy before moving the solution into a production state.

For clarity, the deduplication occurs at the time of dump, and not during the time of the universal share protection point backup.

Tuning universal shares

Two key methods for improving performance of universal shares include tuning the number of vpfsd instances and setting the ingest mode.

Tuning the number of vpfsd instances

A universal share uses one vpfsd instance by default. In most cases, one instance is adequate. Increasing the number of vpfsd instances might improve universal share performance, although it also requires more CPU and memory. You can increase the number of vpfsd instances from 1 to up to 16 and distribute the shares cross all the vpfsd instances.

In extensive testing with a maximum recommended 50 concurrently active shares, setting the number of vpfsd instances to 2 resulted in favorable results without overconsuming CPU and memory resources. Increasing the number of vpfsd

instances to more than 2 should be done carefully and in conjunction with extensive benchmark testing.

There is no requirement to balance the number of universal shares per vpfsd instance because it is done automatically.

To make this change, set the `numOfInstance` value in the `vpfsd_config.json` file on the MSDP server.

Setting the ingest mode

Ingest mode was natively introduced in NetBackup 10.0 to give DBAs a way to increase performance of dumps to the universal share. When ingest mode of a universal share is enabled, the share will persist all the data from memory to disk on the client side at the end of the dump. The ingest mode is faster than normal mode as it does not guarantee all the ingested data is persisted to disk until the ingest mode is turn off. Therefore, turning ingest mode off is critical for data dump integrity.

You can enable and disable the ingest mode for a specific share on the NFS/SMB client side.

To leverage ingest mode, the `.vpfs_special_control_config` file must be updated to turn it on at the beginning of the dump, and then to turn it off at the end of the dump. Ideally, this action would be included in a script that the DBA calls to execute a database dump to the target universal share.

For more information, consult the *NetBackup Deduplication Guide*.

Additional tuning considerations

- Improving performance with emergency engineering binaries (EEBs)
  When dumping MSSQL data to universal shares, performance can be improved in some deployments by applying an EEB. The following table lists the relevant EEBs that are needed for each release of NetBackup.

NetBackup Version	Required EEB
9.1.0.1	4047040 v24
10.0	4070421 v8
10.0.0.1	4078688 v6
10.1	4091734 v6
10.1.1	4102406 v1

- Universal share as a replacement for Oracle incremental forever backup (Co-pilot).

  Beginning in NetBackup 10.1.1, universal share is becoming the replacement for Oracle Co-pilot feature. Internal tests have shown performance can be better with the initial dump, however, the merge part is still bottlenecked with the small I/O updates from Oracle RMAN. This blog describes the Oracle incremental merge feature really well:

  https://blogs.oracle.com/maa/post/using-oracle-incremental-merge

  The merge elapsed time is proportionally increased with the amount of changed data. Internal tests showed that for database change rate greater than 5%, the incremental merge performance can be slower than the full backup. For this reason, we don't recommend the Co-pilot feature if the database change rate is substantially greater than 5%. However, the merge operation has very low overhead on the Oracle client. In other words, normal operation can resume while the merge is in progress with very little impact to Oracle server performance. If this is acceptable, then a universal share can still be a good alternative to stream-based backup even if the change rate is larger.

  Another advantage of the incremental forever backup feature is faster restore, as there is always a latest full image of the database available for access, while with stream-based backups, merging of multiple incremental backup images may be necessary to form the final backup image. Additionally, the Oracle incremental forever backups enables instant access capability that allows you to access the database much faster for certain database operations.

NetBackup for VMware sizing and best practices

Introduction

NetBackup for VMware provides backup and restore of VMware virtual machines (VMs) that run in a vSphere environment. It takes advantage of VMware vSphere Storage APIs – Data Protection (formerly known as VMware vStorage APIs for Data Protection or VADP). NetBackup primarily protects Virtual Machines, backup is configured using 'VMware' as policy type.

Backup operation comprises of different steps that depend on how virtual machines are selected for backup. If VMs are explicitly selected, then one snapshot and one backup job is used per VM to perform the backup. On the other hand, in the case of VMware Intelligent Policy (VIP), first, a discovery job is executed which identifies all the virtual machines matching the criteria set in the policy query, and then one snapshot and one backup job is executed per selected VM. There is only one discovery job for a policy. VIP policies are the recommended way to protect VMware VMs, as they allow use of resource limits to control and limit the load of backup operations on VMware infrastructure. Protection plans are aligned to VIP policies.

Key Components of a NetBackup for VMware setup

- Backup host
  A NetBackup client that performs backups on behalf of the virtual machines. The backup host is the only host on which NetBackup client software is installed. No NetBackup client software is required on the VMware virtual machines. Note that the backup host is referred to as the recovery host when it performs a restore. A NetBackup media server or primary server can also serve as a backup host.
- Discovery host
  Used for the automatic selection of virtual machines for backup via VMware Intelligent Policy (VIP). The discovery host can be on any platform that NetBackup supports for primary or media servers. It can also be the same host as the backup host. A discovery host is configured per policy and defaults to "Backup media server".
- NetBackup media server
  Performs the backups to storage on behalf of the NetBackup client.
- NetBackup primary server
  Manages the backups of virtual machines.

General best practices and notes

- Virtual Machine setup
  When creating virtual machines, use the same name for both hostname and display name. If the policy's Primary VM identifier option is changed, the existing entries on the policy Clients tab still work.
- Simultaneous backups
  VMware recommends that you run no more than four simultaneous backups of virtual machines that reside on the same datastore.
- Backup Host Memory
  Make sure that the VMware backup host has enough memory to handle the number of simultaneous backups that occur.
- VMware backups and MSDP deduplication
  VMware backup to a deduplication storage unit, select the Enable file recovery from VM backup option on the VMware policy tab. This option provides the best deduplication rates. Without the Enable file recovery from VM backup option, the result is a lower rate of deduplication.
- Disabling `Enable file recovery from VM backup`
  If you don't need file-level recoveries you can improve backup performance significantly. This can be especially helpful for a virtual machine with a very large number of small files.

Configuring and controlling NetBackup for VMware

Successful VMware snapshots depend on the amount of I/O that occurs on the virtual machine datastore during the snapshot. When a snapshot is taken, a delta `.vmdk` file is created. The delta disk represents the difference between the running state of the virtual machine and the state that existed at the time of the snapshot. A virtual machine with a large amount of I/Os during snapshot increases the size of the delta disks.

To delete such a snapshot, a large amount of information must be read and written to a disk. This process can reduce the virtual machine performance until the consolidation is complete. The time to delete snapshots and consolidate the snapshot files depends on the amount of data that the guest operating system writes to the virtual disks after you take the last snapshot. The general recommendation is to strive to saturate backup throughput, so the backup duration is as small as possible. Backups should be scheduled when relatively little I/O activity is expected.

Jobs per policy

- VMware intelligent policies (VIPs) - For policies that select virtual machines automatically using the query filter:
  The **Limit jobs per policy** option control the number of parent (discovery) jobs that run simultaneously for the policy. This option does not limit the number of snapshot jobs and backup (`bpbkar`) jobs that the parent job launches.
- VMware policies using manual selection of virtual machines:
  Limit jobs per policy controls the number of virtual machines that the policy can back up simultaneously. Because no discovery job is needed, each virtual machine backup begins with a snapshot job. Each snapshot counts against the Limit jobs per policy setting. If this option is set to 1: the backup of the next virtual machine that is specified in the policy cannot begin until the first snapshot job and its backup are complete.

Limiting NetBackup's load on virtual infrastructure (resource limits)

You can use the **NetBackup Resource Limit** dialog to control the number of simultaneous backups that can be performed on a VMware resource type. The settings apply to all VMware policies. This powerful feature gives you great control over how to maximize performance without causing trashing of virtual infrastructure.

Following are some key resource limits that are available in NetBackup and how to use them to tune the load of backup operations on virtual infrastructure.

Resource limit	Description
vCenter	Maximum jobs that can run per vCenter. Useful to protect a small vCenter server.
Snapshot	Maximum snapshot creates/deletes that can run at the same time. This resource can be useful for limiting the effect that multiple snapshot operations have on the vCenter server. This is a global per-vCenter limit; consider your most resource-constrained vCenter to set the appropriate value. This should be used if NetBackup is reporting failures during snapshot creation or deletion. Snapshot only limits, but will not affect the number of backup jobs.
VMXDatastoret	Controls the maximum number of simultaneous backups per datastore where VMX files are stored. During a backup of a virtual machine snapshot, vmdk writes are cached on this datastore. This resource type is useful for VMs that have VMX and VMDK files distributed across multiple datastores.
Datastore	Maximum number of backup jobs per datastore. Useful to avoid overloading datastore with limited spindles.
ESXserver	Limits jobs running per host.
DatastoreNFSHos	Limits backups per datastore at an NFS Host level.

Discovery

Discovery is the first step when a VMware Intelligent Policy (VIP) is used. This step is applicable only when automatic selection of virtual machines for backup via a VIP is enabled. It discovers virtual machines and filters them using the selection query specified in the policy. The resulting list determines which virtual machines are backed up.

To determine which virtual machines need to be protected, data for all virtual machines for each virtual server (vCenter / ESX) is fetched and evaluated. This is usually a fast operation. Data, once queried, is cached on the discovery host and reused. Typically, data that has been fetched within one hour is reused, but there are circumstances in which new data is fetched. You can control which discovery host is used for each policy; this can have a significant impact on backup and virtual infrastructure.

Consolidated discovery

- Consolidated discovery hosts

In some cases, it might be beneficial to use a fewer number of discovery hosts across policies. Each discovery host will query for all possible virtual machines from every configured virtual machine server by reusing discovery hosts from another policy. Subsequent policies can benefit from the cache built from a previous discovery. Reducing the number of times discovery is done can also reduce the load on the virtual infrastructure. This is especially a problem if various policies with multiple storage units (media servers) share the same backup window.

- Consolidated policies
  Policies are useful to help organize backup lifecycle and provide multiple levels of service as required. For VMware backups, there is one discovery job done for each policy, more number of policies there are more amount of discovery jobs would be running. Consolidating several policies can reduce the number of times vCenters are queried for the data. Careful consideration of all the VMware policies is recommended, especially in the case of a large VMware environment. Another consideration for grouping correct VMs in policies is to have better control over resource utilization across the NetBackup domain.

- Advanced policy attribute - VMware Server List
  This attribute is a mechanism to control which virtual machine servers that NetBackup communicates with for a given policy. In large virtual environments, you can use this list to improve backup/discovery performance. This mechanism can also be used to temporarily avoid a particular vCenter and ESX server, or to eliminate accessing vCenters that are known to not have VMs to be included in the policy.

- DNS Configuration
  NetBackup may be unable to identify virtual machines when you use the **Browse for Virtual Machines** dialog. Virtual machine hostnames may not be properly configured in your Domain Name Server system (DNS), or the DNS system may be slow.
  In a large VMware environment, reverse name lookups can be very slow depending on the number of virtual machines being discovered. You can change the `VNET_OPTIONS` option to determine how many items NetBackup can cache.
  Another option would be to use **VM display name** as the **Primary VM identifier**. This can address problems such as:
  - Host name/DNS not configured correctly
  - DNS server is slow to respond

  However, if using **VM display name** as the **Primary VM identifier**, it is recommended that VM display name is set to same as the VM host name. If **Primary VM identifier** is different than the host name, while doing individual file and folder restores using the Backup, Archive and Restore (BAR) GUI,

backup image selection might not happen automatically. Use the VM search feature to identify correct VM backups to restore from.

Note that each item in NBU's DNS cache takes about 1 KB in memory. This value is in the `bp.conf` file on UNIX and Linux and the registry on Windows. See the NetBackup for VMware Administrator's Guide for details.

Backup and restore operations

Redundancy and performance by using media servers as backup host

Policies can be configured to use the NetBackup media server as a backup host. This allows the performance and throughput of VMware backup operations to scale naturally.

- Storage unit groups
  You can combine the flexibility of backup media servers with a standard feature of NetBackup: storage unit groups. Create a storage unit group that contains the storage units that your media servers can access. Assigning this storage unit as storage for the policy and selecting the **Backup Media Server** as the backup host will allow any of the media servers to operate as a backup host.
  Use storage unit groups with care when you use deduplication storage units, as it might result in a reduction in overall duplication rates. Refer to the NetBackup Deduplication Guide for best practices around storage unit groups.
- Complete LAN-free backups
  Using the correct VMware storage layout and NetBackup storage configuration with a combined Backup Host and Media Server can achieve a complete LAN-free backup.
- Greater Host redundancy
  If one media server goes down, another media server takes over.
- Faster backup
  The media server can read the data from the datastore and send the data straight to the storage device. Without media server access to storage devices, an ordinary backup host must send the backup data over the local network to the media server.

VDDK transport modes

NetBackup support various transport modes supported by underlying vSphere Storage APIs. Each transport mode works the best in specific scenarios. Choosing the right transport mode for an environment is critical for smooth backup or restore operations. Following are some performance, tuning, and compatibility notes of various transport modes and how components of NetBackup interacts with them.

For more detailed notes and configuration steps, refer to the NetBackup for VMware Administrator's Guide

- SAN transport
  The SAN transport mode requires the VMware Backup Host to reside on a physical machine with access to Fibre Channel or iSCSI SAN containing the virtual disks to be accessed. This is an efficient data path because no data needs to be transferred through the production ESX/ESXi host.
  - When using SAN, make sure that datastore LUNs are accessible to the VMware backup host.
  - SAN transport is usually the best choice for backups when running on a physical VMware backup host. However, it is disabled inside virtual machines, so use HotAdd instead on a virtual VMware backup host.
  - SAN transport is not always the best choice for restores. It offers the best performance on thick disks, but the worst performance on thin disks, because of the way vStorage APIs work. For thin disk restore, LAN (NBD) is faster.
  - The job may be slow when you restore to a vCenter Server. For greater speed, designate a VMware Restore ESX server as the destination for the restore.
  - Too many paths. VDDK does not cache device paths. If there are too many paths zoned to a single host, it needs to determine which path to use for each disk, which can slow down the performance. On NetBackup appliances, VxDMP alleviates this problem to some extent, and also load-balances across multiple paths on appliances. However, it is recommended to avoid too many zoned paths.
- NBD transport
  In this mode, the ESX/ESXi host reads data from storage and sends it across a network to the VMware backup host. As its name implies, this transport mode is not LAN-free, unlike SAN transport.
  - The VMware backup server can be a virtual machine, so you can use a resource pool and scheduling capabilities of VMware vSphere to minimize the performance impact of backup. For example, you can put the VMware backup host in a different resource pool than the production ESX/ESXi hosts, with lower priority for backup.
  - Because the data in this case is read by the ESX/ESXi server from storage and then sent to VMware backup host, there must be network connectivity between ESX/ESXi server and VMware backup host. If the VMware backup host has connectivity to the vCenter server but not the ESX/ESXi server, snapshots will succeed but vmdk read/write operations will fail.

- The VMware backup host needs the ability to connect to TCP port 902 on ESX/ESXi hosts while using NBD/NBDSSL for backups and restores.
- VMware uses network file copy (NFC) protocol to read VMDK using NBD transport mode. You need one NFC connection for each VMDK file being backed up. In the older version of vCenter/ESX, there is a limit on the number of NFC connections that can be made per ESX/vCenter server. Backup and restore operations using NBD might hang if this limit is reached. In recent versions of vCenter/ESX, VMware has added additional limits in terms of the total size of NFC connection buffers, which for vSphere 6.x/7.x are 32 or 48 MBs per ESX host.
- Dedicated NICs for NBD. As of vSphere 7.0, ESXi hosts support a dedicated network for NBD transport. This mechanism can be enabled by applying the `vSphereBackupNFC` tag to a NIC using VMware CLI `esxcli`. NetBackup 8.3 and later versions support this.

- HotAdd transport
  When running VMware backup host on a Virtual Machine, vStorage APIs can take advantage of the SCSI HotAdd capability of the ESX/ESXi server to attach the VMDKs of a virtual machine being backed up to the VMware backup host. This is referred to as HotAdd transport mode.
  - HotAdd works only with virtual machines with SCSI disks and is not supported for backing up virtual machines with IDE disks. The paravirtual SCSI controller (PVSCSI) is the recommended default for HotAdd, but other controller types work too.
  - A single SCSI controller can have a maximum of 15 disks attached. To run multiple concurrent jobs totally more than 15 disks it is necessary to add more SCSI controllers to the HotAdd host. A maximum number of 4 SCSI controllers can be added to a HotAdd host, so a total of 60 devices are supported at the maximum.
  - HotAdd requires that the VMware backup host have access to datastores where the virtual machine being backed up resides. This essentially means:
    - ESX where the VMware backup host is running should have access to datastores where the virtual machine being backed up resides.
    - Both the VMware backup host and virtual machine being backed up should be under the same datacenter.
  - HotAdd cannot be used if the VMFS block size of the datastore containing the virtual machine folder for the target virtual machine does not match the VMFS block size of the datastore containing the VMware backup host virtual machine. For example, if you back up a virtual disk on a datastore with 1MB

blocks, the VMware backup host must also be on a datastore with 1MB blocks.

Transport mode notes:

- NBD transport
  If the attempt to restore a full virtual machine fails while using the SAN transport type, try the NBD transport type instead.
- Slow performance of NBD
  Restoring a virtual machine with a transport mode of NBD or NBDSSL may be slow in the following cases:
  - The virtual machine had many small data extents due to heavy fragmentation. (A file system extent is a contiguous storage area defined by block offset and size.)
  - The restore is from a block-level incremental backup and the changed blocks on the disk were heavily fragmented when the incremental backup occurred.

  For faster restores in either of these cases, use the hotadd transport mode instead of NBD or NBDSSL.

For more details and troubleshooting, see the following Veritas Knowledge Base articles:

- VMware Transport Modes: Best practices and troubleshooting
- Two causes of slow NetBackup for VMware restore performance

Best practices: Storage lifecycle policies (SLPs)

This topic provides tuning best practices for storage lifecycle policies (SLPs).

Data flow and SLP design best practices

When constructing Storage Lifecycle Policies (SLP), it is important to construct a data flow that considers several factors resulting in a sustainable solution. These may include, but are not limited to:

- Typical backup windows and duration
- Backup deduplication rates
- Image sizes (many small images less than 5 GB, few large images greater than 10 TB, etc.)
- Workload types (database, stream handlers, NDMP, VMware, etc.)
- Source and target storage types (Data Domain, MSDP, cloud, tape, etc.)

- Retention levels
- Immutability duration
- Encryption requirements
- Total number of desired copies
- Network bandwidth
- Performance/throughput expectations
- Concurrent read/write IO to the source and destination storage

Definition of a data flow

A data flow refers to the paths the data takes between copies. For example, with optimized duplication or replication with OpenStorage (OST), deduplicated data on the source is cached and then unique segments are transferred to the target. With most workload types and deduplication rates in the 90% range this is very efficient with little IO impact.

A second example is OST to tape. This data is rehydrated on the source storage server into its raw native format, transferred, and then written to tape storage. This data flow is much more IO-intensive than optimized duplication or replication and is best done as a final copy and as sparingly as possible. When an SLP is configured, the following items are defined: the number of copies, retention of each copy, the storage target for each copy, and the order in which each copy is made. The creation of these copies can be scheduled but is managed automatically by the `nbstserv` service in NetBackup.

Types of storage targets

A storage target can be located on a BYO server, a Veritas appliance, a 3rd party OST target, a private or public cloud storage target, or tape. When choosing a storage target it is important to choose the most performant storage target for the backup in order to complete the primary backup copy as quickly as possible. High-performance storage targets like MSDP residing upon storage HW that provides the highest level of speed and resiliency should be prioritized for the primary backup copy.

It is also important to determine what performance features are supported with the storage target, like deduplication, compression, and I/O bandwidth (that is, the speed of each spindle and number of spindles). When using MSDP as your primary storage target for backup, replications and duplications to MSDP and MSDP-C leverage optimized duplication, which minimizes storage utilization and optimizes duplication speed.

Creating an SLP that copies data from dissimilar OST devices (such as MSDP to Data Domain) is not recommended. Rehydration and an additional deduplication

cycle will create performance bottlenecks that significantly reduce performance and also reverses the benefits of other features. For example, a 3rd-party OST storage target uses a different deduplication engine than MSDP/MSDP-C. Therefore, a copy of the backup image from MSDP/MSDP-C to/from a 3rd party OST storage target will require full backup image rehydration and then full image deduplication. These steps take longer to complete and consume more memory, CPU, network, and I/O resources than would be required when leveraging the optimized duplication feature. Basically, once a backup image is deduplicated, any subsequent copies made to different storage targets do not require rehydration and another deduplication activity as long as the OST devices are similar (MSDP to MSDP/C, Data Domain to Data Domain, etc.).

For long-term retention (LTR) copies, the use of tape may be chosen. Although we support duplication from OST to tape, a full rehydration of the data is required. This operation is resource intensive and reduces performance significantly when done at scale.

When using public cloud storage, it is important to consider both the performance costs and the financial impacts. Performance factors to consider include the speed, resiliency, and sizing of the storage tier that was chosen for the copy, as well as the network bandwidth between the source storage and target storage. Financial impacts are also extremely important, such as deduplication rates, total bucket size, retention targets, and the cost of warming and restoring the data. Optimized duplication will reduce the total data transferred compared to traditional cloud (cloud without deduplication).

The use of storage unit groups (STUGs) is not recommended with SLP other than with tape storage units. There are two main reasons for this:

- NetBackup evaluates the sum of all the streams of the disk pools, the sum of all active jobs using all the storage units for the disk pools, and then has to add/subtract those from SLP workgroups. This is very resource intensive when done at scale and leads to significant delays in job submissions and excessive queueing that impacts backup operations as well. This is less significant for tape because the concurrent jobs that are written to a tape storage unit is relatively low (typically dozens, potentially 100) compared to the concurrent jobs to a disk storage server (typically hundreds, potentially thousands).
- Optimized duplication is not supported for disk storage units in STUG, even with similar storage device types. The data is rehydrated, transferred in full, then deduplicated again at the target where this is used, and incurs all of the negative impacts of doing so. Immutability is not supported with STUG configurations, with or without SLP.

Sizing a storage target

When designing an SLP configuration, it is important to properly size the backup copy storage so that most restore operations will be done from it for the best performance. From a resource perspective it is more expensive to restore from secondary copies such as cloud or tape than from a local copy or an on-premises copy. Longer retention copies require more space and needs to be considered for secondary copy storage.

Number of copies and retention

When determining the number of copies to be retained of each backup image, it is best not to apply a single universal approach to all data. For example, less critical images such as QA/dev should not require as many copies or be retained as long as critical production images. Treating them the same leads to extraneous copies, space, network consumption, and other impacts. Consider prioritizing or tiering the data into separate SLPs to remove inefficiency.

SLP operations should cascade with retentions, where copy 1 is the shortest retention for high performance, copy 2+ is longer than copy 1, and copy 3 is the longest. Creating more than three copies with SLP in a large configuration is not a best practice, as the resource overhead needed to create and track the additional copies with a large volume of them compounds exponentially from a database perspective.

Immutability objectives

Objectives including immutability should be considered, and implemented where WORM duration increases with each subsequent copy, similar to backup retention. If copy 1 is locked for too long it may cause capacity management problems and impact backup operations. Data cannot be easily deleted to make space for backups in this scenario. If an SLP backlog starts to build, this may impact the expiration and cleanup of WORM images also.

Data characteristics

Data characteristics are attributes of specific data such as segment object size, encryption (MSDP, 3rd-party, KMS, etc.), compression, database with stream handlers, VMs with stream handlers, etc. These operations are often most critical for backup copies.

How backup copies are created impacts secondary copies. Take the following scenarios for example:

- Images with exponentially more segments (DB stream handler) have much larger caches, require more memory, and/or may take longer to process if the cache limit for them is exceeded.

- Data compressed or encrypted before the source backup is taken leads to low deduplication rates, often zero. It is best to use native features such as Data in-Transit Encryption (DTE), MSDP encryption, KMS, Client Direct, etc., to have NetBackup deduplicate the data first and then encrypt it.
- Individual large images such as over 10 TB may require much larger cache sizes and take significantly longer to restore, duplicate, or rehydrate. Increasing the cache for these increases memory use for the storage servers involved and must be considered in designing an SLP.
- If third-party encryption must be used on the source data, consider not deduplicating it and instead writing to non-OST storage to avoid the performance impacts.

Not properly considering the data characteristics for an SLP configuration often leads to poor performance, large SLP backlogs, and capacity management problems. The simplest configurations are often the best performing and most scalable configurations. Carefully consider the requirements end to end for all copies and the resources for each to avoid large-scale problems in an environment.

Network bandwidth

When designing an SLP configuration, consider the network path that the data will traverse. Keep in mind that the ideal configuration has copy 1 designed for the fastest backup and restore and shortest retentions, where copies 2+ are longer retention copies.

Traversing a slow or unstable network, excessive rehydration, shared network links between sites or applications, and other external factors such as bandwidth limits can all have significant impacts on the SLP configuration.

Order of storage target priority

As a best practice, the fastest and most resilient backup storage should be targeted as the primary location for the initial backup and most of the anticipated restores. Each additional copy should have the same or increasing retention period with the longest term retention (LTR) copy being the final one. Often these reside in the cloud, tape, or Access appliances.

SLP tuning

NetBackup provides a number of tunables for SLP processing such as:

- SLP windows to defer or schedule secondary operations.
- Window close behaviors that cancel or stop submitting jobs after a certain time.
- Batching logic:
  - FIFO: First In, First Out

    - LIFO: Last In, First Out
- Minimum and maximum batch sizes
- Resource multipliers (controls how many jobs can be submitted simultaneously)

There are more parameters outlined in the NetBackup Administrator's Guide, Volume I for the desired NetBackup version.

Key points

The entire environment must be considered to design an efficient, sustainable SLP configuration. Think end-to-end what is required to create each copy, the path it must take, the resources required to create each, and the impacts of doing so.

Broadly speaking, the best practices include:

- Tier the data to protect the most critical first and the less critical second.
- The simplest configurations are usually the best performing ones.
- When using OST, do not use dissimilar devices in the same SLP.
- Avoid rehydration and duplication to tape as much as possible.
- Use native NetBackup features for encryption, compression, stream handlers, etc., over third party features when possible.
- Create copy 1 on the fastest storage available, and size it to satisfy most of the anticipated resource requests.
- Create subsequent copies with the same or longer retention in sequence where the final copy has the longest retention in the SLP.

Targeted SLP

A large number of A.I.R. Replication jobs running concurrently can cause performance problems due to contention. Limiting the number of concurrent A.I.R. Replications can allow each of the active replication jobs to perform at their maximum potential given the underlying environment. At NetBackup 8.3, it is possible to limit the number of concurrent A.I.R. Replication jobs by setting a `SLP.REPLICATION_TARGET_JOB_LIMIT` value. This value can be applied globally, affecting all A.I.R. Replications, or by target storage server. This setting limits the number of A.I.R. Replications going to a target storage server.

`SLP.REPLICATION_TARGET_JOB_LIMIT = <limit_spec>[,<limit_spec>][,...]`

It is possible to throttle replication jobs globally or per target server. Each *limit-spec* value can be a numeric value which sets the limit for every target storage server. It can also take the form *<storage_server_name>:<number>* which sets the limit for a specific target storage server. The two limit types can be specified on the same

parameter setting, as shown in the following examples. (First enter `nbsetconfig` from a command line. Then enter the `SLP.REPLICATION_TARGET_JOB_LIMIT` values.)

```
# Set the replication limit for each target storage server to 10.
->nbsetconfig
nbsetconfig>SLP.REPLICATION_TARGET_JOB_LIMIT = 10
nbsetconfig><end-file marker - Unix: Ctrl+D Enter, Windows: Ctrl+Z Enter>

# Set the limit for two named target storage servers to 12 and 6.  Set the limit for
# all other target storage servers to 8.
->nbsetconfig
nbsetconfig>SLP.REPLICATION_TARGET_JOB_LIMIT= targetServerA:12, targetServerB:6, 8
nbsetconfig><end-file marker - Unix: Ctrl+D Enter, Windows: Ctrl+Z Enter>
```

The advantage of Target A.I.R. is that it helps to gain more granular control over the number of Replication jobs that will run at one time to specific storage servers acting as targets. Prior to this, the only option was to greatly increase the `Maximum size per A.I.R. replication job` to every increasing levels, but this would not set a specific number of running operations. Also, different target storage servers were all treated the same, no considerations for the capabilities of different storage servers were possible. By adding the `SLP.REPLICATION_TARGET_JOB_LIMIT` parameter, limits can be set on the number of A.I.R. Replications that can run per target, while still keeping the `Maximum size per A.I.R. replication job` parameters in place to match needs. Also, specific limits can be put on specified storage servers to accommodate the capabilities and conditions of different storage servers.

The `SLP.REPLICATION_TARGET_JOB_LIMIT` parameter is limited to A.I.R. Replications. More information is available in the following technical article:

How to tune NetBackup Auto Image Replication (A.I.R.) operations for maximum performance

Replicating small images can create delays between the replication of each image. Between each image there is a check done to make sure the image is replicated. With small images this check can be longer than the time needed to actually replicate the image which can add up to a significant time. It is possible to adjust this delay for checking the image using `AIR_POLL_INTERVAL_TIME` setting. The size of the image to check can also be adjusted using the `AIR_POLL_INTERVAL_CHUNK_SIZE` setting.

- `AIR_POLL_INTERVAL_TIME = <seconds>`

  This option sets the time factor to a fixed number of seconds and should be used when it's desirable to optimize the time needed to complete the replication. It will generate the most overhead. Use this option with an initial value of 1 to

minimize the time needed and increase the throughput. Increase the value if the overhead generated impacts system performance.

- `AIR_POLL_INTERVAL_CHUNK_SIZE = <size_value>`
  This option attempts to calculate the best time factor based on image size. The goal is to set the time factor to approximate the time it will take to replicate a typical image. The size value should be a decimal number representing the number of bytes that can be replicated in 1 second. The default value is 104857600 (100 megabytes). Increasing the chunk size value will decrease the time factor and improve the accuracy of the completion detection at the cost of increased overhead.

Limiting the number of SLP secondary operations to maximize performance

Less is more. Run fewer jobs but allow each job to manage as much data as possible. SLP parameters can be used to assist the lifecycle to run duplication jobs more efficiently. To prevent the lifecycle from running numerous small duplication jobs in frequent succession, NetBackup accumulates lists of similar images into a batch. Then, each batch of images is copied as a set in one duplication job, instead of one image at a time. The NetBackup administrator can change how large the batch file scan becomes, or how frequently batch jobs are requested.

More information is available in the following technical articles:

Storage Lifecycle Policy (SLP) tuneable parameters to optimize the duplication process

Storage Lifecycle Policy (SLP) Cheat Sheet

Also see *SLP Parameters properties* in the NetBackup Administrator's Guide, Volume I

Table 6-4 Common SLP parameters that impact BID files

SLP property	Default value
`Maximum size per duplication job` The largest batch size that can run as a single duplication job.	100 GB
`Minimum size per duplication job` The smallest batch size that can run as a single duplication job. The job does not run until enough images accumulate to reach this minimum batch size or until the Force interval for small jobs time is reached.	8 GB

Table 6-4 Common SLP parameters that impact BID files *(continued)*

SLP property	Default value
`Maximum size per A.I.R. replication job` The largest batch size that can run as a single job for Auto Image Replication.	100 GB
`Maximum Images per A.I.R. Import job` The largest number of images in a single batch that can run as an Auto Image Replication import job.	250
`Minimum Images per A.I.R. Import job` The fewest number of images in a single batch that can run as an Auto Image Replication import job. The job does not run until either the minimum size is reached or the Force interval for small jobs time is reached.	1
`Force interval for small job` The age that the oldest image in a batch must reach after which the batch is submitted as a duplication job. This value prevents many small duplication jobs from running at one time or running too frequently. It also prevents NetBackup from waiting too long before it submits a small job.	30 minutes

Effects of increasing these SLP parameters:

- `Minimum size per duplication job` and `Force interval for small job`
  Increasing the `Minimum size per duplication job` will cause the total number of images in a BID file to reach a higher size threshold before a small job will be started. The small job will be started after the `Force interval for small job` value is reached. This can have the effect of duplication jobs starting further apart because it will take more time for a total batch size to reach the limit.
- `Maximum size per duplication job`
  Increasing the `Maximum size per duplication job` will cause the maximum batch size for a single duplication job be raised to the specified size. A higher value can have the effect of reducing the number of duplication jobs that can run.
- `Maximum size per A.I.R. replication job`
  Increasing the `Maximum size per A.I.R. replication job` will cause the maximum batch size for a single A.I.R. replication job be raised to the specified size. A higher value can have the effect of reducing the number of A.I.R. replication jobs that run while managing more images at one time.

- `Maximum Images per A.I.R. Import job`

  Increasing the `Maximum Images per A.I.R. Import job` will cause the maximum number of images in a single job in an A.I.R. Import. A higher value can have the effect of reducing the number of A.I.R. Import jobs that run, while managing more images per job.

- `Minimum Images per A.I.R. Import job`

  Increasing the `Minimum Images per A.I.R. Import job` will cause the minimum number of images in a single job be increased and then wait until the `Force interval for small job` is reached. The effect of increasing this value is that fewer A.I.R. Import jobs run, but more images be managed with each job.

Storage Server IO

Storage unit `Concurrent Jobs` value and the disk pool `Max IO Streams per volume` value:

- The `Max IO Streams` value for the disk pool sets the number of read/write operations that can run at the same time on the disk pool. The `Concurrent Jobs` value on the storage unit sets the maximum backup/restore jobs that can run at the same time to the storage unit. More than one storage unit can point to the same disk pool.
- It is important that the total number of `Concurrent Jobs` from storage units that point to the same disk pool must not exceed the configured `Max IO Streams` value for the disk pool. So that secondary operations can perform, the total number of Concurrent Jobs set in the storage units that point to the same disk pool should not be higher than 90% of the total of the `Max IO Streams` value of the disk pool. This 10% allows the possibility of secondary operations running while backup/restore operations are maximized. Setting `REPLICATION_TARGET_JOB_LIMIT` to set limits on the number of A.I.R. Replication per target storage server helps to maximize the performance of each secondary operation. SLP parameters can be used to limit the number of secondary operations.

In summary, a disk pool is configured by default to have an unlimited value for `Max IO Streams`. It is recommended that the `Max IO Streams` for the pool be limited to a level that maximizes the number of jobs that can run, but is not too high that the disk pool performance is depreciated due to too many read/write operations occurring at the same time.

A storage unit that is configured to point to the disk pool should have a `Concurrent Jobs` value lower than the `Max IO Streams` value for the disk pool. The storage unit's `Concurrent Jobs` value will address backup jobs and restore jobs, but does

not impact the number of secondary operations like duplications, replications, import type jobs. These secondary operations are managed by the SLPs and try to run as many as possible, usually dictated by the `Max IO Streams` count.

More information about configuring disk pools is also available:

See "Best practices: Disk pool configuration - setting concurrent jobs and maximum I/O streams" on page 103.

Best practices: NetBackup NAS-Data-Protection (D-NAS)

Refer to the following guides on the NetBackup Documentation Landing Page for more details about NetBackup for NAS-Data-Protection (D-NAS):

- *NetBackup NAS Administrator's Guide*

Best practices: NetBackup for Nutanix AHV

Refer to the following guides on the NetBackup Documentation Landing Page for more details about NetBackup for Nutanix AHV:

- *NetBackup Web UI Nutanix AHV Administrator's Guide*
- *NetBackup for Nutanix Acropolis Hypervisor (AHV) Administrator's Guide*

Tuning parameters - resource throttling

To avoid the load on the Nutanix AHV production environment, set resource limit options for AHV backups:

- The backup job resource limits can be set at the AHV cluster, per storage container and per AHV host.
- The snapshot job resource limit can be set at, AHV cluster level.
- Set the global level resource throttling to apply across clusters, containers, and AHV host level. Consider the load on AHV to decide on these restrictions.
- Override the global level throttling values if you need to alter those values specific to given cluster or container or AHV host level.

Tuning parameters - Media server load balancing

- Media server load balancing is dependent on the storage server used. It is supported only for the Media Server Deduplication Pool and Cloud storage.

- For the Media server Load Balancing feature, discovery would be done only by one of the available media servers even though the backup jobs are distributed across multiple media servers.
- For Nutanix AHV backups using the Media server Load Balancing option, it is required that all media servers assigned to a particular storage are of the same version.

Best practices: NetBackup Sybase database

Database memory allocation

During installation and upgrade processes, NetBackup interrogates the total system memory and allocates between 1 GB and 1/3 of the total system memory to the Sybase database server. This allocation can be adjusted by increasing the value of the `-ch` option in `server.conf`.

The `server.conf` file is found in these locations:

- Windows:
  `<install_path>\VERITAS\NetBackupDB\conf\server.conf`
- UNIX:
  `/usr/openv/var/global/server.conf`

More information about the `-ch` option is available:

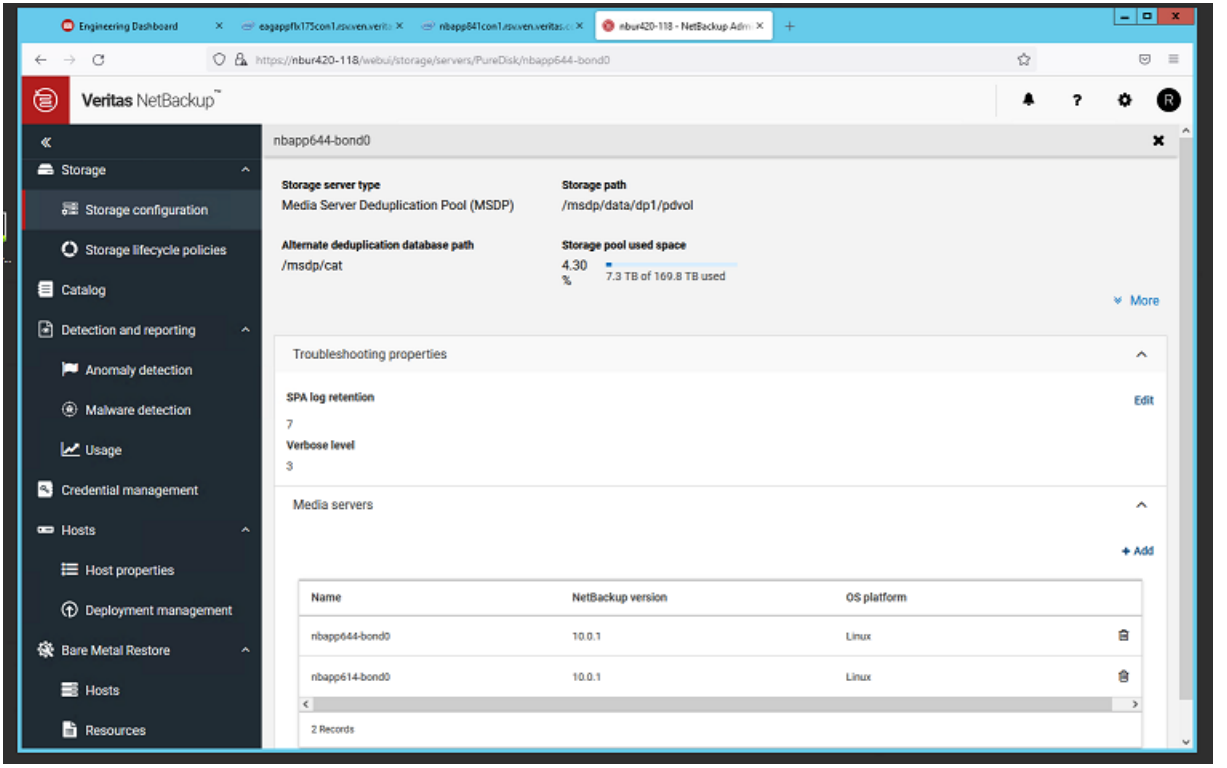
-ch Database Server Option (SAP SQL Anywhere Server - Database Administration)

Best practices: Avoiding media server resource bottlenecks with Oracle VLDB backups

Very Large Database (VLDB) support is a NetBackup 10.1 feature that enables Oracle database backups to span across multiple MSDP storage units or multiple nodes in a cluster for better performance.

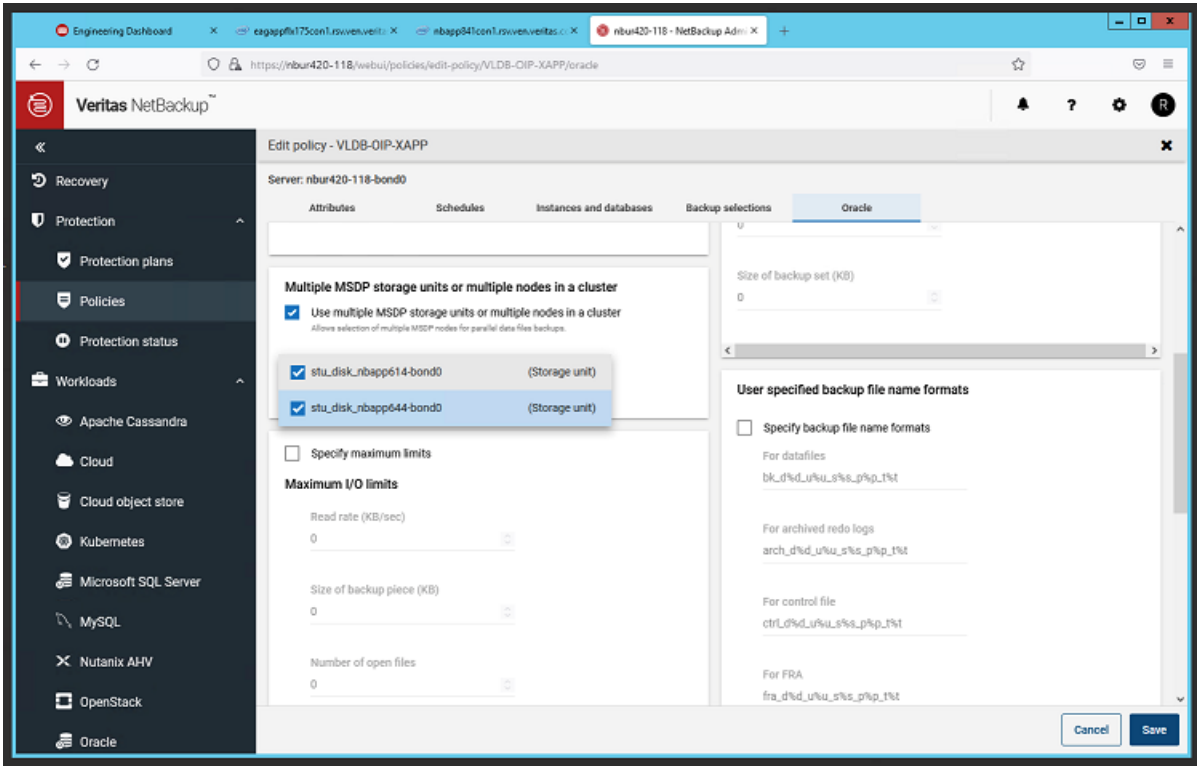
VLDB support is designed to remedy media server resource bottlenecks by redirecting workloads to one or more MSDP storage units or cluster nodes.

When you configure MSDP storage units in the NetBackup web UI, it is critical to add all defined media servers for optimal load balancing and maximum performance, as shown in this screen shot:



Then in the NetBackup web UI:

1. Create an Oracle type policy.
2. Populate all fields as you do for a standard Oracle agent.
3. In the **Oracle** tab, select multiple MSDP storage units or cluster nodes.
4. Save the policy as in this screen shot:



Best practices: Avoiding media server resource bottlenecks with MSDPLB+ prefix policy

Beginning in NetBackup 10.1, a new feature to address the media/storage server bottleneck is available for NetBackup servers configured with scale-out NetBackup, such as NBFS (NetBackup Flex Scale) or Veritas Cloud Scale Technology. While Very Large Database (VLDB) supports only Oracle database backups, the MSDPLB+ feature works for all policies. It can help balance workloads (Oracle, MS-SQL, and so on) across the available cluster nodes. The MSDPLB+ feature can be enabled by adding the prefix *MSDPLB+* to the policy name.

Performance of both features have been optimized in the release of NetBackup 10.1.1 with the Global Deduplication feature implementation. Based on internal Oracle workload testing and analysis, in some configuration, MSDPLB+ can outperform VLDB. Veritas recommends using MSDPLB+, if available, to ensure balanced workload distribution across the cluster nodes.

Best practices: Cloud deployment considerations

Review the following considerations before you deploy cloud resources with NetBackup:

- Virtual machine sizing
  When selecting the virtual machine size follow the steps that are discussed in the chapters for sizing primary servers (Chapter 2) and media servers (Chapter 3). These recommendations also apply to deployments in the cloud.
- NetBackup Installation
  NetBackup installation options are available that use predefined systems in the cloud provider marketplace. Veritas provides these systems to the cloud vendors, and they include options for different NetBackup versions. Current customers may use their existing license key (BYOL); otherwise, contact Veritas Sales to request a key. If you chose to use another VM, then make sure that the installation disk is large enough for the NetBackup installation, with room for the NetBackup logs.
- Operating systems
  For NetBackup to operate optimally, you may need to change or tune various operating system settings, for example firewall settings, number of open files, and so on. Refer to the *NetBackup Installation Guide* for these recommendations.
- Disk cache for cloud upload and download
  The NetBackup cloud-tier allows each media server to create one or more cloud LSU. It is important to know that for each cloud LSU created, roughly 1 TB of MSDP pool is reserved for the LSU to be used as cloud disk cache. For a MSDP pool with limited storage size, the reserved disk cache can consume too much space, resulting in little usable space for regular backup jobs. If jobs are failing with error code 129 and 84, it may indicate that there is no space left on the device, even though the MSDP pool may still have plenty of available space according to `df -h` and `dsstat`. This is the default behavior. However, if the system has enough memory, and the `useMemForUpload`is set to true, the disk cache may not not be needed for upload. In this case, the `CloudUploadCacheSize` in `contentrouter.cfg`can be lowered to bypass the issue.

Chapter 7

Measuring Performance

This chapter includes the following topics:

Measuring NetBackup performance: overview

The final measure of NetBackup performance is the following:

- The length of time that is required for backup operations to complete (usually known as the backup window).
- The length of time that is required for a critical restore operation to complete.

However, to measure and improve performance calls for performance metrics more reliable and reproducible than wall clock time. This chapter discusses these metrics in more detail.

After establishing accurate metrics as described here, you can measure the current performance of NetBackup and your system components to compile a baseline performance benchmark. With a baseline, you can apply changes in a controlled

way. By measuring performance after each change, you can accurately measure the effect of each change on NetBackup performance.

How to control system variables for consistent testing conditions

For reliable performance evaluation, eliminate as many unpredictable variables as possible to create a consistent backup environment. Only a consistent environment can produce reliable and reproducible performance measurements. Note that it is essential to ensure that the performance of the backup environment is reproducible. Without it, you won't know if the performance difference is due to tuning or just the run-to-run variation.

This topic explains some of the variables to consider as they relate to the NetBackup server, the network, the NetBackup client, or the data itself.

Table 7-1 System variables to control for testing

Variables	Considerations for controlling
Server variables	Eliminate all other NetBackup activity from your environment when you measure the performance of a particular NetBackup operation. Areas to consider include the automatic scheduling of backup jobs by the NetBackup scheduler, and background tasks that wake up periodically, such as CRQP (Content Router Queue Processing), which by default wakes up twice every day at 00:20 AM and 12:20 PM. CRQP, compaction, and CRC check may negatively impact the measured performance When policies are created, they are usually set up to allow the NetBackup scheduler to initiate the backups. The NetBackup scheduler initiates backups according to the following: traditional NetBackup frequency-based scheduling, or on certain days of the week, month, or other time interval. This process is called calendar-based scheduling. As part of the backup policy, the **Start Window** indicates when the NetBackup scheduler can start backups using either frequency-based or calendar-based scheduling. When you perform backups to test performance, this scheduling might kick in and interfere with the performance test. The NetBackup scheduler may initiate backups unexpectedly, especially if the operations you intend to measure run for an extended period of time. See “Running a performance test without interference from other jobs” on page 146.

Table 7-1 System variables to control for testing *(continued)*

Variables	Considerations for controlling
Network variables	Network performance is key to optimum performance with NetBackup. Ideally, you should use a separate network for testing, to prevent unrelated network activity from skewing the results. In many cases, a separate network is not available. If not, ensure that non-NetBackup activity is kept to a minimum during the test. If possible, schedule the test when backups are not active. Even occasional or sudden increases of network activity may be enough to skew the test results. If you share the network with production backups occurring for other systems, you must account for this activity during the test. Another network variable is host name resolution. NetBackup depends heavily upon a timely resolution of host names to operate correctly. If you have any delays in host name resolution, try to eliminate that delay. An example of such a delay is a reverse name lookup to identify a server name from an incoming connection from an IP address. You can use the `HOSTS` (Windows) or `/etc/hosts` (Linux/UNIX) file for host name resolution on systems in your test environment.
Client variables	Make sure that the client system is relatively quiescent during performance testing. A lot of activity, especially disk-intensive activity such as Windows virus scanning, can limit the data transfer rate and skew the test results. Do not allow another NetBackup server, such as a production server, to access the client during the test. NetBackup may attempt to back up the same client to two different servers at the same time. The results can be severely affected for a performance test that is in progress. Different file systems have different performance characteristics. It may not be valid to compare data throughput on Linux/UNIX VxFS or Windows FAT file systems to Linux/UNIX NFS or Windows NTFS systems. In addition, different OS releases may introduce changes that can also affect system performance. For such a comparison, factor the difference between the OS releases and the file systems into your performance tests and into any conclusions.

Table 7-1 System variables to control for testing *(continued)*

Variables	Considerations for controlling
Data variables	Monitoring the data you back up improves the repeatability of performance testing. If possible, move the data you use for testing to its own drive or logical partition (not a mirrored drive). Defragment the drive before you begin performance testing. For testing restores, start with an empty disk drive or a recently defragmented disk drive with ample empty space. For testing backups to tape, always start each test with an empty piece of media, as follows: ■ Expire existing images for that piece of media through the Catalog node of the NetBackup Administration Console, or run the `bpexpdate` command. ■ Another approach is to use the `bpmedia` command to freeze any media containing existing backup images so that NetBackup selects new media for the backup operation. This step reduces the effect of backup data location placement which can effect backup performance and yields more consistent results between tests. It also reduces mounting and unmounting of the media that has NetBackup catalog images and that cannot be used for normal backups. When you test restores, always restore from the same backup image on the media to achieve consistent results between tests. A large set of data generates a more reliable, reproducible test than a small set of data. Startup and shutdown overhead within the NetBackup operation will probably skew a performance test with a small data set. These variables are difficult to keep consistent between test runs and are therefore likely to produce inconsistent test results. A large set of data minimizes the effect of startup and shutdown times. Design the data set to represent the makeup of the data in the intended production environment. If the data set in the production environment contains many small files on file servers, the data set for the tests should also contain many small files. A representative data set can more accurately predict the NetBackup performance that can be expected in a production environment. The type of data can help reveal bottlenecks in the system. Files that contain non-compressible (random) data increase the amount of I/O against the storage media and affect the backup performance, particularly when the data has low deduplication ratio. As long as the other components of the data transfer path can keep up, you may find the storage media is the bottleneck. On the other hand, files containing highly-compressible data can be processed at higher rates when hardware compression is enabled. The result may be a higher overall throughput and may expose the network as the bottleneck. Many values in NetBackup provide data amounts in kilobytes and rates in kilobytes per second. For greater accuracy, divide by 1024 rather than rounding off to 1000 when you convert from kilobytes to megabytes or kilobytes per second to megabytes per second.

Running a performance test without interference from other jobs

Use the following procedure to run a performance test. This procedure helps prevent the NetBackup scheduler from running other backups during the test.

To run a performance test

1. Create a policy specifically for performance testing.
2. Leave the schedule's **Start Window** field blank.

   This policy prevents the NetBackup scheduler from initiating any backups automatically for that policy.
3. To prevent the NetBackup scheduler from running backup jobs unrelated to the performance test, consider setting all other backup policies to inactive.

   You can use the **Deactivate** command from the NetBackup Administration Console. You must reactivate the policies after the test, when you want to start running backups again.
4. Before you start the performance test, check the Activity Monitor to make sure no NetBackup jobs are in progress.
5. To gather more logging information, set the legacy and unified logging levels higher and create the appropriate legacy logging directories.

   By default, NetBackup logging is set to a minimum level. Note that higher log levels may reduce performance, depending on the tests and the log levels.

   For details on how to use NetBackup logging, refer to the logging chapter of the NetBackup Logging Reference Guide. Keep in mind that higher logging levels consume more disk space and can affect NetBackup performance. The best practice for performance testing is to keep the default log level setting initially and increase the log level only if need to troubleshooting a specific performance issue.
6. From the policy you created for testing, run a backup on demand.

   Click **Actions > Manual Backup** in the NetBackup Administration Console.

   Or, you can use a user-directed backup to run the performance test. However, the **Manual Backup** option is preferred. With a manual backup, the policy contains the entire definition of the backup job. The policy includes the clients and files that are part of the performance test. If you run the backup manually from the policy, you can be certain which policy is used for the backup. This approach makes it easier to change and test individual backup settings, from the policy dialog.

7 During the performance test, check for non-NetBackup activity on the server and try to reduce or eliminate it.

8 Check the NetBackup Activity Monitor after the performance test for any unexpected activity that may have occurred during the test, such as a restore job.

About evaluating NetBackup performance

You can obtain statistics on NetBackup data throughput from these tools:

- The NetBackup Activity Monitor
- The NetBackup All Log Entries report

Select the reporting tool according to the type of NetBackup operation you want to measure:

- Non-multiplexed backup (disk and tape)
- Multiplexed backup (tape only)
- Restore

Table 7-2 Where to obtain NetBackup performance statistics

Operation to report on	In Activity Monitor	In All Log Entries report
Non-multiplexed backup	Yes	Yes
Multiplexed backup	No (see next column)	Yes Obtain the overall statistics from the All Log Entries report. Wait until all the individual backup operations which are part of the multiplexed backup are complete. In this case, the statistics available in the Activity Monitor for each of the individual backup operations are relative only to that operation. The statistics do not reflect the total data throughput to the tape drive.
Restore	Yes	Yes

The statistics from these two tools may differ, because of differences in rounding techniques in the Activity Monitor versus the All Logs report. For a given type of operation, choose one of the tools and consistently record your statistics only from that tool. In both the Activity Monitor and the All Logs report, the data-streaming speed is reported in kilobytes per second. If a backup or restore is repeated, the reported speed can vary between repetitions depending on many factors. Factors

include the availability of system resources and system utilization. The reported speed can be used to assess the performance of the data-streaming process.

The statistics from the NetBackup error logs show the actual amount of time reading and writing data to and from the storage device. The statistics do not include the time for mounting and positioning the tape. You should cross-reference the information from the error logs with data from the bpbkar log on the NetBackup client. (The bpbkar log shows the end-to-end elapsed time of the entire process.) Such cross references can indicate how much time was spent on operations unrelated to reading and writing to the storage device.

Evaluating NetBackup performance through the Activity Monitor

To evaluate performance through the NetBackup Activity Monitor

1. Run the backup or restore job.
2. Open the NetBackup Activity Monitor.

3 Verify that the backup or restore completed successfully.

The Status column should contain a zero (0).

4 View the log details for the job by selecting the **Actions > Details** menu option, or by double-clicking on the entry for the job. Then select the **Detailed Status** tab.

Obtain the NetBackup performance statistics from the following fields in the Activity Monitor:

Start Time/End Time	These fields show the time window during which the backup or restore job took place.
Elapsed Time	This field shows the total elapsed time from when the job was initiated to job completion. It can be used as an indication of total wall clock time for the operation.
KB per Second	The data throughput rate.
Kilobytes	Compare this value to the amount of data. Although it should be comparable, the NetBackup data amount is slightly higher because of administrative information (metadata) that is saved for the backed-up data.
Deduplication rate	Pay close attention to the deduplication rate for each job. If there are significant numbers of jobs running with very poor deduplication rates, it can lead to overall poor backup performance. If possible consider grouping clients with poor deduplication rate and running them in a separate backup window.

For example, if you display properties for a directory that contains 500 files, each 1 megabyte in size, the directory shows a size of 500 megabytes. (500 megabytes is 524,288,000 bytes, or 512,000 kilobytes.) The NetBackup report may show 513,255 kilobytes written, reporting 1255 kilobytes more than the file size of the directory. This report is true for a flat directory. Subdirectory structures may diverge due to the way the operating system tracks used and available space on the disk.

Note that the operating system may report how much space was allocated for the files in question, not only how much data is present. If the allocation block size is 1 kilobyte, 1000 1-kilobyte files report a total size of 1 megabyte, although only 1 kilobyte of data exists. The greater the number of files, the larger this discrepancy may become.

Evaluating NetBackup performance through the All Log Entries report

To evaluate performance through the All Log Entries report

1. Run the backup or restore job.
2. Run the All Log Entries report from the NetBackup reports node in the NetBackup Administrative Console. Be sure that the Date/Time Range that you select covers the time period during which the job was run.
3. Verify that the job completed successfully by searching for entries such as the following:

   For a backup: "the requested operation was successfully completed"

   For a restore: "successfully read (restore) backup ID..."
4. Obtain the NetBackup performance statistics from the messages in the report.

Table of NetBackup All Log Entries report

Table 7-3 describes messages from the All Log Entries report.

The messages vary according to the locale setting of the primary server.

Table 7-3 Messages in All Log Entries report

Entry	Statistic
`started backup job for client <name>, policy <name>, schedule <name> on storage unit <name>`	The **Date/Time** column for this entry show the time at which the backup job started.
`successfully wrote backup id <name>, copy <number>, <number> Kbytes`	For a multiplexed backup, this entry shows the size of the individual backup job. The **Date/Time** column indicates when the job finished writing to the storage device. The overall statistics for the multiplexed backup group are found in a subsequent entry. These statistics include the data throughput rate to the storage device,
`successfully wrote <number> of <number> multiplexed backups, total Kbytes <number> at Kbytes/sec`	For multiplexed backups, this entry shows the overall statistics for the multiplexed backup group including the data throughput rate.

Table 7-3 Messages in All Log Entries report *(continued)*

Entry	Statistic
`begin writing backup id <name>, copy <number>, fragment <number>, destination path <name>`	The **Date/Time** column for this entry shows when the restore job began reading from the storage device.
`successfully wrote backup id <name>, copy <number>, fragment <number>, Kbytes at <number> Kbytes/sec`	For non-multiplexed backups, this entry combines the information in the previous two entries for multiplexed backups. The single entry shows the following: ■ The size of the backup job ■ The data throughput rate ■ When the job finished writing to the storage device (in the **Date/Time** column).
`the requested operation was successfully completed`	The **Date/Time** column for this entry shows the time at which the backup job completed. This value is later than the "successfully wrote" entry (in a previous message): it includes extra processing time at the end of the job for tasks such as NetBackup image validation.
`begin reading backup id <name>, (restore), copy <number>, fragment <number> from media id <name> on drive index <number>`	The **Date/Time** column for this entry shows when the restore job started reading from the storage device. (Note that the latter part of the entry is not shown for restores from disk, because it does not apply.)
`successfully restored from backup id <name>, copy <number>, <number> Kbytes`	For a multiplexed restore, this entry shows the size of the individual restore job. (As a rule, all restores from tape are multiplexed restores, because non-multiplexed restores require additional action from the user.) The **Date/Time** column indicates when the job finished reading from the storage device. The overall statistics for the multiplexed restore group are found in a subsequent entry below. These statistics include the data throughput rate.
`successfully restored <number> of <number> requests <name>, read total of <number> Kbytes at <number> Kbytes/sec`	For multiplexed restores, this entry shows the overall statistics for the multiplexed restore group, including the data throughput rate.

Table 7-3 Messages in All Log Entries report *(continued)*

Entry	Statistic
`successfully read` (restore) `backup id <name>, copy <number>, fragment <number>, <number> Kbytes at <number> Kbytes/sec`	For non-multiplexed restores, this entry combines the information from the previous two entries for multiplexed restores. The single entry shows the following: ■ The size of the restore job ■ The data throughput rate ■ When the job finished reading from the storage device (in the **Date/Time** column) As a rule, only restores from disk are treated as non-multiplexed restores.

Additional information on the NetBackup All Log Entries report

For other NetBackup operations, the NetBackup All Log Entries report has entries that are similar to those in the following table:

See Table 7-3 on page 150.

For example, it has entries for image duplication operations that create additional copies of a backup image. The entries may be useful for analyzing the performance of NetBackup.

The `bptm` debug log file for tape backups (or `bpdm` log file for disk backups) contains the entries that are in the All Log Entries report. The log also has additional detail about the operation that may be useful. One example is the message on intermediate data throughput rate for multiplexed backups:

```
... intermediate after number successful, number Kbytes at
number Kbytes/sec
```

This message is generated whenever an individual backup job completes that is part of a multiplexed backup group. For example, the debug log file for a multiplexed backup group (that consists of three individual backup jobs) may include the following: two intermediate status lines, then the final (overall) throughput rate.

For a backup operation, the `bpbkar` debug log file also contains additional detail about the operation that may be useful.

Note that writing the debug log files during the NetBackup operation introduces overhead that may not be present in a production environment. Factor that additional overhead into your calculations.

The information in the All Logs report is also found in the following locations:

- Linux/UNIX
  `/usr/openv/netbackup/db/error`

- Windows

  install_path\NetBackup\db\error

See the NetBackup Logging Reference Guideto learn how to set up NetBackup to write these debug log files.

Evaluating system components

In addition to your evaluation of NetBackup's performance, you should also verify that common system resources are in adequate supply.

On Windows: For high-level information, you can use the Windows Task Manager. For more detailed information, use the Windows Performance Monitor utility.

See "About the Windows Performance Monitor" on page 159.

For further information on the Windows Performance Monitor, refer to your Microsoft documentation.

For Linux/Unix: Multiple commands are available to monitor system resources depending on the resource you want to monitor. Most commonly used system monitoring commands are: `sar`, `vmstat`, `iostat`, and `ifstat`.

About measuring performance independent of tape or disk output

You can measure the disk (read) component of NetBackup's speed independent of the network components and tape components.

The following techniques are available:

- `bpbkar`
  This technique is easier.
- SKIP_DISK_WRITES touch file
  This technique may be helpful in more limited circumstances.
- NULL storage units
  See the following tech note for more information:
  How to configure a "null" Storage Unit for NetBackup 8.x and 9.x Media servers

Note: In these procedures, the primary server is the client.

Measuring performance with bpbkar

Use this procedure to measure disk performance with `bpbkar`.

To measure disk I/O using the bpbkar command

1. Turn on the legacy `bpbkar` log by ensuring that the `bpbkar` directory exists:
   - Linux/UNIX:

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

   - Windows:

     ```
     install_path\NetBackup\logs\bpbkar
     ```

2. Set the logging level to 1.
3. Enter the following:
   - Linux/UNIX:

     ```
     /usr/openv/netbackup/bin/bpbkar -nocont -dt 0 -nofileinfo
       -nokeepalives file system > /dev/null
     ```

     Where *file system* is the path being backed up.
   - Windows:

     ```
     install_path\NetBackup\bin\bpbkar32 -nocont X:\ > NUL
     ```

     Where `X:\` is the path being backed up.
4. Check how long it took NetBackup to move the data from the client disk:
   - Linux/UNIX:
     The start time is shown in the `bpbkar initialize` entry in the `bpbkar` log. The end time is shown in the entry `bpbkar main: INF - Client completed sending data for backup`. The elapse time is the time duration between the start time and the end time.
     The start time is the first `PrintFile` entry in the `bpbkar` log. The end time is the entry "Client completed sending data for backup." The amount of data is given in the entry "Total Size."
   - Windows:
     Check the `bpbkar` log for the entry "Elapsed time". For example:
     `Elapsed time: 1622 secs 59370573 bps`

Bypassing disk performance with the SKIP_DISK_WRITES touch file

The `SKIP_DISK_WRITES` procedure can be used on Linux/UNIX or Windows.

The SKIP_DISK_WRITES procedure is a useful follow-on to the bpbkar procedure. The bpbkar procedure may show that the disk read performance is not the bottleneck. If it is not the bottleneck, the bottleneck is in the data transfer between the client bpbkar process and the server bptm process. The following SKIP_DISK_WRITESprocedure may be helpful.

If the SKIP_DISK_WRITES procedure shows poor performance, the problem may involve the network, or shared memory (such as not enough buffers, or buffers that are too small). You can change shared memory settings.

See "About shared memory (number and size of data buffers)" on page 177.

Warning: The following procedure can lead to data loss. The SKIP_DISK_WRITES touch file disables all backup data write operations for disk. It is not recommended to touch this file on a production server or data lost will occur. Disable active production policies for the duration of this test. You must remove the touch file when this test is complete. (See step 7.)

To bypass disk I/O using the SKIP_DISK_WRITES touch file

1. Create a new disk storage unit, with /tmp or some other directory as the image directory path.
2. Create a policy that uses the new disk storage unit.
3. Deactivate any active production policies for the duration of this test.
4. Create the SKIP_DISK_WRITES file.

   Linux/UNIX:

   ```
   /usr/openv/netbackup/db/config/SKIP_DISK_WRITES
   ```

   Windows:

   ```
   install_path\Netbackup\db\config\SKIP_DISK_WRITES
   ```

   This file disables all data-write operations for disk backups but retains the creation of disk fragments and associated metadata.
5. Run a backup from this policy.

   NetBackup creates a file in the storage unit directory as if this backup is a real backup to disk. The image file is 0 bytes long.

6 To remove the zero-length file and clear the NetBackup catalog of a backup that cannot be restored, run this command:

Linux/UNIX:

```
/usr/openv/netbackup/bin/admincmd/bpexpdate -backupid backupid -d 0
```

Windows:

```
install_path\Netbackup\bin\admincmd\bpexpdate -backupid backupid -d 0
```

Where *backupid* is the name of the file that resides in the storage unit directory.

7 Remove the `SKIP_DISK_WRITES` file.

8 Re-activate any policies that were deactivated for this procedure.

Measuring performance with the GEN_DATA directive (Linux/UNIX)

You can use the `GEN_DATA` directive to test I/O speed between `bpbkar` and the storage unit. For further information, see the following Veritas tech note:

How to use the GEN_DATA file list directives with NetBackup for UNIX/Linux Clients for Performance Tuning

Monitoring Linux/UNIX CPU load

1 Use the `vmstat` utility to monitor CPU usage.

2 In the `vmstat` output: Add the **us** and **sy** CPU columns to get the total percentage of the CPU that is being used for user tasks and system tasks. Consistently non-zero **wa** indicates there are processes waiting for IO. You need to check the disk service time with `iostat` when seeing consistently high **wa** values.

Note: The `sar`, `mpstat`, and `top` commands are useful tools for monitoring CPU usage

---

Monitoring Linux/UNIX memory use

1 Use the `vmstat` utility to memory usage.

2 In the `vmstat` output: Add free, buffer and cache columns to estimate the available physical memory size. The **si** and **so** columns indicate the amount of swapping activity taking place, consistent non-zero **si** and **so** values in the hundreds or thousands indicate physical memory shortage.

Monitoring Linux/UNIX disk load

You can use the `iostat` utility to check device I/O performance, such as average wait time and percentage of disk utilization.

Try the following:

```
iostat -ktxN 5
```

where 5 specifies a five-second refresh rate.

Sample output from the Red Hat 7 kernel:

```
iostat -ktxN 5
Time: 07:39:14 AM
avg-cpu:  %user   %nice %system %iowait  %steal   %idle
           5.02    0.00    8.84    0.84    0.00   85.30

Device:  rrqm/s  wrqm/s   r/s     w/s    rkB/s   wkB/s avgrq-sz avgqu-sz   await  svctm  %util
sda        0.00     .40    0.00    3.80    0.00    48.80    25.68     0.11   29.89   3.37   1.28
sdb        0.00    7.60    0.80    9.40    3.20    69.60    14.27     0.52   50.90   3.53   3.60
sys-r      0.00    0.00    0.00    2.00    0.00     8.00     8.00     0.02   10.80   2.40   0.48
sys-swap   0.00    0.00    0.00    0.00    0.00     0.00     0.00     0.00    0.00   0.00   0.00
sys-usr    0.00    0.00    0.00    0.00    0.00     0.00     0.00     0.00    0.00   0.00   0.00
sys-h      0.00    0.00    0.80   27.20    3.20   108.80     8.00     1.33   47.63   1.11   3.12
sdc        0.00    2.60    0.00    3.60    0.00    25.60    14.22     0.01    3.56   0.22   0.08
sdd        0.00    0.00    0.00    0.00    0.00     0.00     0.00     0.00    0.00   0.00   0.00
```

Note: This example is from a Red Hat 7 Linux system. Other operating systems may use different options. Refer to the `iostat` man page of each OS for details.

Helpful report values are the following:

- `await`: The average time (in milliseconds) for device I/O requests to complete, for both the virtual device and the physical disk. This includes the time the requests spend waiting in the disk queue and the time servicing them. In general, low average wait values indicate better throughput.
- `%util`: The percentage of elapsed time in which I/O requests were sent to the device. As the value approaches 100%, saturation of the device occurs. A lower percentage is better.

If NetBackup is running on Red Hat 8 Linux and the Veritas storage manager, InfoScale, is used for MSDP storage management, you may find the VxVM devices name no longer appear in the `iostat` output. You can check the VxVM IO mode with the command `vxtune vol_use_rq`. If the current value of `vol_use_rq` is '0', then BIO mode is enabled. Otherwise, Request mode is enabled. The mode change

is needed to bypass a known Red Hat 8 Linux bug. In order to analyze IO statistics for VxVM devices when BIO mode is enabled, use `vxstat` instead of `iostat`.

Monitoring Linux/UNIX network traffic

The Linux `sar` utility can be used to monitor network traffic.

The sample output below is generated from the following command:

```
sar -n DEV 5
```

where 5 specifies a five second refresh rate.

Sample output:

```
07:05:35 AM     IFACE   rxpck/s   txpck/s    rxkB/s    txkB/s   rxcmp/s   txcmp/s  rxmcst/s
07:05:40 AM     eth13      0.00  19100.20      0.00   1231.12      0.00      0.00      0.00
07:05:40 AM     eth10   8503.20      1.00 280410.05      0.06      0.00      0.00      0.00
07:05:40 AM     eth12  12662.80  11310.60 392599.48    733.85      0.00      0.00      0.00
07:05:40 AM     bond0  92385.60  80866.80 2330765.89   5659.79      0.00      0.00      0.00
07:05:40 AM      eth0      0.00      0.00      0.00      0.00      0.00      0.00      0.00
07:05:40 AM      eth1      1.00      0.00      0.11      0.00      0.00      0.00      0.00
07:05:40 AM      eth2      0.00      0.00      0.00      0.00      0.00      0.00      0.00
07:05:40 AM      eth3      0.00      0.00      0.00      0.00      0.00      0.00      0.00
07:05:40 AM      eth4   9381.20  30225.00 312068.98   1948.09      0.00      0.00      0.00
07:05:40 AM      eth5  22071.80   3743.20 483838.00    368.05      0.00      0.00      0.00
07:05:40 AM      eth6      0.00   7364.40      0.00    476.56      0.00      0.00      0.00
07:05:40 AM      eth7      0.00   1004.00      0.00    378.48      0.00      0.00      0.00
07:05:40 AM      eth8      0.00      1.00      0.00      0.06      0.00      0.00      0.00
07:05:40 AM      eth9   6372.60   1277.80 204399.06     82.35      0.00      0.00      0.00
07:05:40 AM        lo   9210.60   9210.60  65180.86  65180.86      0.00      0.00      0.00
07:05:40 AM     eth11  33394.40   6839.80 657449.78    441.19      0.00      0.00      0.00
```

If a port is having fixed ingress/egress throughput and the throughput column in `rxKB/s` or `txKB/s` is close to the bandwidth of the network port, the system may be network bound. See the chapter *Tuning the NetBackup data transfer path* for information about how to increase network bandwidth with bonding

Monitoring Linux/Unix system resource usage with dstat

The commands `vmstat`, `iostat`, and `netstat` are useful for monitoring the resource usage of each hardware component separately. If you want to monitoring all four major resources (CPU, memory, I/O, and network) concurrently, you would need to open multiple windows, one window for each command.

`dstat` is a versatile tool that you can use to replace all three commands mentioned above. The output from `dstat` without any option can display the usage pattern of the four major hardware resources in the same report side by side. For example, the following example is the output from running command, `dstat 5` while running 8 concurrent backup streams on a 5340 NetBackup Appliance:

```
----total-cpu-usage---- -dsk/total- -net/total- ---paging-- ---system--
usr sys idl wai hiq siq|  read writ| recv  send|    in out |  int  csw
 13   9  77   0   0   1| 438k 3405M|1767M 3242k|     0   0 | 163k 126k
 13   8  77   0   0   1| 362k 3476M|1836M 3157k|     0   0 | 160k 131k
 13   8  78   0   0   1| 421k 3391M|1822M 3390k|     0   0 | 154k 127k
 13   8  77   0   0   1| 523k 3319M|1666M 2843k|     0   0 | 166k 121k
 14   8  77   0   0   1| 587k 3443M|1833M 3399k|     0   0 | 171k 127k
 13   8  78   0   0   1| 539k 3483M|1796M 3374k|     0   0 | 149k 117k
 12   8  79   0   0   1| 389k 3178M|1692M 3484k|     0   0 | 153k 117k
 10   7  83   0   0   1|1256k 2574M|1269M 2707k|     0   0 | 147k 113k
 10   7  82   0   0   1| 306k 2571M|1417M 2938k|     0   0 | 137k 115k
 10   6  83   0   0   1| 405k 2500M|1226M 3154k|     0   0 | 120k 109k
 10   6  83   0   0   1| 397k 2606M|1341M 3073k|     0   0 | 128k 110k
```

For system monitoring and performance troubleshooting, being able to see all four major resources usage pattern side by side can help speeding up root cause analysis.

About the Windows Performance Monitor

The Performance Monitor organizes information by object, counter, and instance.

An object is a system resource category, such as a processor or physical disk. Properties of an object are counters. Counters for the **Processor** object include **%Processor Time**, which is the default counter, and **Interrupts/sec**. Duplicate counters are handled by instances. For example, to monitor the **%Processor Time** of a specific CPU on a multiple CPU system, the **Processor** object is selected. Then the **%Processor Time** counter for that object is selected, followed by the specific CPU instance for the counter.

In the Performance Monitor, you can view data in real-time format or collect the data in a log for future analysis. Specific components to evaluate include CPU load, memory use, and disk load.

Note: You should use a remote host for monitoring of the test host to reduce any load that might otherwise skew results.

Monitoring Windows CPU load

Use the following procedure to determine if the system has enough power to accomplish the requested tasks.

To monitor Windows CPU load

1. Start the Windows Performance Monitor.

   For instructions, refer to your Microsoft documentation.

2. To determine how hard the CPU is working, monitor the **% Processor Time** counter for the **Processor** object.

   For **% Processor Time**, values of 0 to 80 percent are generally safe. Values from 80% to 90% indicate that the system is heavily loaded. Consistent values over 90 percent indicate that the CPU is a bottleneck.

   Spikes close to 100 percent are normal and do not necessarily indicate a bottleneck. However, if sustained loads close to 100 percent are observed, consider tuning the system to decrease process load, or upgrade to a faster processor.

3. To determine how many processes are actively waiting for the processor, monitor the **Process Queue Length** counter for the **System** object.

   Sustained **Processor Queue Lengths** greater than 2 indicate that too many threads are waiting to be executed. To correctly monitor the **Processor Queue Length** counter, the Performance Monitor must track a thread-related counter. If you consistently see a queue length of 0, verify that a non-zero value can be displayed.

   The default scale for the **Processor Queue Length** may not be equal to 1. Be sure to read the data correctly. For example, if the default scale is 10x, then a reading of 40 means that only 4 processes are waiting.

Monitoring Windows memory use

Memory is a critical resource for increasing the performance of backup operations.

To monitor Windows memory use

1. Start the Windows Performance Monitor.

   For instructions, refer to your Microsoft documentation.

2. To examine memory usage, view information on the following:

Committed Bytes	Committed Bytes displays the size of virtual memory that has been committed, as opposed to reserved. Committed memory must have disk storage available or must not require the disk storage because the main memory is large enough. If the number of Committed Bytes approaches or exceeds the amount of physical memory, you may encounter issues with page swapping.
Page Faults/sec	Page Faults/sec is a count of the page faults in the processor. A page fault occurs when a process refers to a virtual memory page that is not in its working set in main memory. A high Page Fault rate may indicate insufficient memory.

Monitoring Windows disk load

To use disk performance counters to monitor the disk performance in Performance Monitor, you may need to enable the counters. Windows may not have enabled the disk performance counters by default for your system.

To get more information about disk performance counters

- Enter the following:

```
diskperf -help
```

To enable the counters and allow disk monitoring

1. Enter the following:

```
diskperf -y
```

2. Restart the system.

To disable the counters and cancel disk monitoring

1. Enter the following:

```
diskperf -n
```

2. Restart the system.

To monitor disk performance

1. Use the **%Disk Time** counter for the **PhysicalDisk** object.

   Track the percentage of elapsed time that the selected disk drive is servicing read or write requests.

2. Monitor the **Avg. Disk Queue Length** counter and watch for values greater than 1 that last for more than one second.

   Values greater than 1 for more than a second indicate that multiple processes are waiting for the disk to service their requests.

See "Bypassing disk performance with the SKIP_DISK_WRITES touch file" on page 154.

Increasing disk performance

You can use the following techniques to increase disk performance:

- Refer to the following topic on measuring disk performance.
  See "Bypassing disk performance with the SKIP_DISK_WRITES touch file" on page 154.
- Check the fragmentation level of the data.
  A highly fragmented disk limits throughput levels. Use a disk maintenance utility to defragment the disk.
- Consider adding additional disks to the system to increase performance.
  Multiple processes writing to the same disk/LUN create I/O contention, which can significantly slowing IO performance. Adding more disks/LUNs to reduce the I/O contention can significantly improve the overall system performance, especially if the workload is I/O intensive, such as the high number of concurrent backup streams with a relatively low deduplication ratio or a system with concurrent backup, restore, duplication or replication jobs running concurrently.
- Determine if the data transfer involves a compressed disk.
  Windows drive compression adds overhead to disk read or write operations, with adverse effects on NetBackup performance. Use Windows compression only if it is needed to avoid a disk full condition.
- Consider converting to a system with a Redundant Array of Independent Disks (RAID).
  Though more expensive, RAID devices offer greater throughput and (depending on the RAID level) improved reliability.
- Determine what type of controller technology drives the disk.
  A different system might yield better results.

Chapter 8

Tuning the NetBackup data transfer path

This chapter includes the following topics:

About the NetBackup data transfer path

The overall performance of NetBackup is limited by the slowest component in the backup system. For example, a fast tape drive that is combined with an overloaded server yields poor performance. A fast tape drive on a slow network also yields poor performance.

The backup system is referred to as the data transfer path. The path usually starts at the data on the disk and ends with a backup copy on tape or disk.

This chapter subdivides the standard NetBackup data transfer path into four components: the NetBackup client, the network, the NetBackup server, and the storage device.

This chapter discusses NetBackup performance evaluation and improvement from a testing perspective. It describes ways to isolate performance variables to learn

the effect of each variable on overall system performance. It also describes how to optimize NetBackup performance with regard to those variables. It may not be possible to optimize every variable on your production system.

Note: The requirements for database backups may not be the same as for file system backups. This information in this chapter applies to file system backups unless otherwise noted.

About tuning the data transfer path

This chapter contains information on ways to optimize NetBackup. This chapter is not intended to provide tuning advice for particular systems. For help fine-tuning your NetBackup installation, please contact Veritas Consulting Services.

Before trying particular tuning steps, consider the following:

- Ensure that your system meets NetBackup's recommended minimum requirements
  Refer to the *NetBackup Installation Guide* and *NetBackup Release Notes* for information about these requirements. For better performance, follow the best practices as specified in Chapter 2 and 3.
- Ensure that you have the most recent NetBackup software patch installed
- Know your hardware
  Many performance issues can be traced to hardware or other environmental issues. You must understand the entire data transfer path to determine the maximum obtainable performance in your environment. Poor performance is often the result of poor planning, which results from unrealistic expectations of components of the transfer path.
  More information about NetBackup server configuration is available:
  See "NetBackup hardware design and tuning considerations" on page 44.

Tuning suggestions for the NetBackup data transfer path

In every backup system there is room for improvement. To obtain the best performance from a backup infrastructure is not complex, but it requires careful review of the many factors that can affect processing. The first step is to gain an accurate assessment of each hardware component and networking component in the backup data path. Many performance problems may be caused by inadequate

hardware configuration and can be resolved by adjusting the hardware before attempting to change NetBackup parameters.

NetBackup software offers many resources to help isolate performance problems and assess the effect of configuration changes. However, it is essential to thoroughly test both backup and restore processes after making any changes to the NetBackup configuration parameters.

This topic provides practical ideas to improve your backup system performance and avoid bottlenecks.

You can find background details in the following NetBackup manuals:

NetBackup Administrator's Guide, Volume I

NetBackup Troubleshooting Guide

Table 8-1 Tuning suggestions for the NetBackup data path

Tuning suggestions	Description
Use multiplexing	Multiplexing writes multiple data streams from several clients to a single tape drive or several tape drives. Multiplexing can improve the backup performance of slow clients, multiple slow networks, and many small backups (such as incremental backups). Multiplexing reduces the time each job waits for a device to become available. It thereby makes the best use of the transfer rate of your storage devices. See "How fragment size affects restore of a multiplexed image on tape" on page 206. Refer also to the NetBackup Administrator's Guide, Volume II for more information about using multiplexing.
Stripe a disk volume across drives.	A striped set of disks can pull data from all drives concurrently, to allow faster data transfers.
Maximize the use of your backup windows	You can configure all your incremental backups to happen at the same time every day. You can also stagger the execution of your full backups across multiple days. Large systems can be backed up over the weekend while smaller systems are spread over the week. You can start full backups earlier than the incremental backups. They might finish before the incremental backups and return all or most of your backup window to finish the incremental backups.
Convert large clients to SAN Clients	A SAN Client is a client that is backed up over a SAN connection to a media server rather than over a LAN. SAN Client technology is for large databases and application servers where large data files are rapidly read from disk and streamed across the SAN. SAN Client is not suitable for file servers where the disk read speed is relatively slow. See "Best practices: NetBackup SAN Client" on page 98.

Table 8-1 Tuning suggestions for the NetBackup data path *(continued)*

Tuning suggestions	Description
Use network bonding to join two or more network interfaces together to form a single interface	Network bonding simplifies network management and offers expanded network bandwidth and performance improvement over a single interface. It also improves the network redundancy: when one interface is down or unplugged, the other interfaces in the bonding can still work.
Avoid a concentration of servers on one network	If many large servers back up over the same network, convert some of them to media servers or attach them to private backup networks. Either approach decreases backup times and reduces network traffic for your other backups.
Use a dedicated media server for NetBackup operations	For a backup server, use a dedicated media server for backups only. Using a server that also runs several applications unrelated to backups can severely affect your performance and maintenance windows.
Consider the requirements of backing up your catalog	Remember that the NetBackup catalog needs to be backed up. To facilitate NetBackup catalog recovery, the primary server should have access to a disk storage server, a cloud server, or dedicated tape drive, either stand-alone or within a robotic library.
Level the backup load	To improve multi-stream backup performance for tape backups, you can use multiple drives and spread the load across them. Similarly, for multi-stream disk backups, you can configure multiple file systems on disks/LUNs. For best disk I/O performance, avoid multiple file systems from sharing the same disk or LUN.
Consider bandwidth limiting	Bandwidth limiting lets you restrict the network bandwidth that is consumed by one or more NetBackup clients on a network. The bandwidth setting appears under **Host Properties > Primary Servers, Properties**. The actual limiting occurs on the client side of the backup connection. This feature only restricts bandwidth during backups. Restores are unaffected. When a backup starts, NetBackup reads the bandwidth limit configuration and then determines the appropriate bandwidth value and passes it to the client. As the number of active backups increases or decreases on a subnet, NetBackup dynamically adjusts the bandwidth limiting on that subnet. If additional backups are started, the NetBackup server instructs the other NetBackup clients that run on that subnet to decrease their bandwidth setting. Similarly, bandwidth per client is increased if the number of clients decreases. Changes to the bandwidth value occur on a periodic basis rather than as backups stop and start. This characteristic can reduce the number of bandwidth value changes.

Table 8-1 Tuning suggestions for the NetBackup data path *(continued)*

Tuning suggestions	Description
Try throttling at different levels	NetBackup provides ways to throttle loads between servers, clients, policies, and devices. Note that these settings may interact with each other: compensating for one issue can cause another. The best approach is to use the defaults unless you anticipate or encounter an issue. Try one or more of the following: ■ Adjust the backup load on the server. Change the **Limit jobs per policy** attribute for one or more of the policies that the server backs up. For example, you can decrease **Limit jobs per policy** to reduce the load on a server on a specific subnetwork. Reconfigure policies or schedules to use storage units on other servers. Use bandwidth limiting on one or more clients. ■ Adjust the backup load on the server during specific time periods only. Reconfigure schedules to use storage units on the servers that can handle the load (if you use media servers). ■ Adjust the backup load on the clients. Change the **Maximum jobs per client** global attribute. An increase to **Maximum jobs per client** can increase the number of concurrent jobs that any one client can process and therefore increase the load. ■ Reduce the time to back up clients. Increase the number of jobs that clients can perform concurrently, or use multiplexing. Increase the number of jobs that the server can perform concurrently for the policies that back up the clients. ■ Give preference to a policy. Increase the **Limit jobs per policy** attribute value for the preferred policy relative to other policies. Alternatively, increase the priority for the policy. ■ Adjust the load between fast and slow networks. Increase the values of **Limit jobs per policy** and **Maximum jobs per client** for the policies and clients on a faster network. Decrease these values for slower networks. Another solution is to use bandwidth limiting. ■ Limit the backup load that one or more clients produce. Use bandwidth limiting to reduce the bandwidth that the clients use. ■ Maximize the use of devices Use multiplexing for tape devices or multiple file systems for disk devices. Also, allow as many concurrent jobs per storage unit, policy, and client as possible without causing server, client, or network performance issues. ■ Prevent backups from monopolizing devices. Limit the number of devices that NetBackup can use concurrently for each policy or limit the number of drives per storage unit. Another approach is to exclude some of the devices from Media Manager control.

See “NetBackup client performance in the data transfer path” on page 168.

See “NetBackup network performance in the data transfer path” on page 170.

See “NetBackup server performance in the data transfer path” on page 176.

See “NetBackup storage device performance in the data transfer path” on page 210.

NetBackup client performance in the data transfer path

Many factors can affect the NetBackup client component of the NetBackup data transfer path. Consider the following to identify possible changes that may improve NetBackup performance:NetBackup

Table 8-2 Factors that affect the client component of the NetBackup data transfer path

Factors	Notes
Disk fragmentation	Fragmentation severely affects the data transfer rate from the disk. Fragmentation can be repaired using disk management utility software.
Number of disks	Add disks to the system to increase performance. If multiple processes attempt to log data simultaneously, divide the data among multiple physical disks.
Disk arrays	Convert to a system that is based on a Redundant Array of Inexpensive Disks (RAID). RAID devices generally offer greater throughput and (depending on the RAID level) improves reliability.
SAN client	For critical data that requires high bandwidth for backups, consider the SAN client feature. Refer to the SAN Client and Fibre Transport Guide.
Type of controller technology that drives the disk	Different controller technology could yield different results.
Virus scanning	Virus scanning can severely affect the performance of the NetBackup client, especially for systems such as large Windows file servers. Consider disabling virus scans during backup or restore.

Table 8-2 Factors that affect the client component of the NetBackup data transfer path *(continued)*

Factors	Notes
NetBackup notify scripts	The NetBackup notify scripts are very useful in certain situations, such as shutting down a running application to back up its data. However, these scripts must be written with care to avoid any unnecessary lengthy delays at the start or end of the backup. If the scripts do not perform tasks essential to the backup, remove them. Linux/UNIX notify scripts: ■ `bpstart_notify` ■ `bpend_notify` Windows notify scripts: ■ `bpstart_notify.bat` ■ `bpend_notify.bat`
NetBackup software location	If the data being backed up is on the same physical disk as the NetBackup installation, note: performance may be adversely affected, especially if NetBackup debug log files are generated. If logs are used, the extent of the degradation is greatly influenced by the logs' verbose setting. If possible, install NetBackup on a separate physical disk to avoid disk drive contention.
Snapshots (hardware or software)	If snapshots are taken before the backup of data, the time that is needed to take the snapshot can affect the performance.
NetBackup Client Job Tracker	If the Job Tracker is running on the client, NetBackup estimates the data to be backed up before the backup. Gathering this estimate affects the startup time and the data throughput rate, because no data is written to the NetBackup server during this estimation. **Note:** The Job Tracker is disabled by default. If it is launched, it runs until the user logs out. Avoid running the NetBackup Client Job Tracker if the data-gathering process takes too long.
Determine the theoretical performance of the NetBackup client software	Use the NetBackup client command `bpbkar` (Linux/UNIX) or `bpbkar32` (Windows) to determine how fast the NetBackup client can read the data to be backed up. You may be able to eliminate data read speed as a performance bottleneck. See "About measuring performance independent of tape or disk output" on page 153.

NetBackup network performance in the data transfer path

To improve the overall performance of NetBackup, consider the following network components and factors.

Network interface settings

Make sure that your network connections are properly installed and configured.

Note the following:

- Network interface cards (NICs) for NetBackup servers and clients must be set to full-duplex. Do not use Auto-sense or Auto-negotiate.
- Both ends of each network cable (the NIC card and the switch) must be set identically as to speed and mode. (Both NIC and switch must be at full duplex.) Otherwise, link down, excessive or late collisions, and errors result.
- If auto-negotiate is used, make sure that both ends of the connection are set at the same mode and speed.
- In addition to NICs and switches, all routers must be set to full duplex.
- Using AUTOSENSE may cause network problems and performance issues.
- Consult the operating system documentation for instructions on how to determine and change the NIC settings.
- If multiple switches are needed to meet the network need for a single NetBackup domain, methods are available to combine multiple switches including stacking, trunking, and uplinks.
- Consider network bonding to join multiple network interfaces into a single interface to increase network bandwidth on a single interface and improve redundancy.
- For 10 gigabit or faster network connections, consider implementing jumbo frames. Be aware that jumbo frames must be supported through all devices in the data transfer paths. Consult with your networking team about the use of jumbo frames in your environment.

Network load

To evaluate remote backup performance, consider the following:

- The amount of network traffic
- The amount of time that network traffic is high

Small bursts of high network traffic for short durations can decrease data throughput rate. However, if the network traffic remains high, in particular if the MB/s in or MB/s out is frequently close to the network interface bandwidth, for example, 1.2GB/sec on a 10Gbps link, the network is probably the bottleneck. Try to schedule backups when network traffic is low. If your network is loaded, you may want to implement a secondary network which is dedicated to backup and restore traffic.

Alternatively, you may add additional network interfaces and use network bonding to join multiple network interfaces together to form a single interface to increase the single interface bandwidth.

Note also: to check the network, use FTP to transfer a large file (50 gigabytes or more) from the media server to the client and back again. Observe how long each operation takes. If moving the file in either direction takes significantly longer than the other, the network has a problem.

Setting the network buffer size for the NetBackup media server

A modern, well configured, operating system with properly written TCP drivers is unlikely to need TCP memory tuning by NetBackup. Accordingly, the best NetBackup configuration is to disable tuning by placing a zero (0) into the `NET_BUFFER_SZ` file on media servers and Linux/UNIX clients. Simply deleting the file is not equivalent because some NetBackup processes have default `setsockopt` API calls configured to overcome past external problems with various platforms and drivers. See the following paragraphs for more details and other considerations.

`NET_BUFFER_SZ` is a tunable media server parameter that can be used to adjust the size of the network buffer space. The operating system uses this buffer space for the connection between the `bptm` child process and the client process. This buffer space caches either received data from the network (a backup) or written data to the network (a restore). The parameter sets the value for NetBackup to use for the network buffer space, but the operating system may not allow the change.

The NetBackup media server can be configured to request that the operating system use a non-default size for the network buffer space. If the `NET_BUFFER_SZ` touch file exists, `bptm` requests that the operating system adjust the size. The operating system may or may not allow the change, depending on the operating system revision and the current TCP tuning.

The following examples are from `bptm` logs on various platforms. These examples show how `bptm` records the size that was used and any previous size requested by NetBackup.

For example:

Red Hat

```
setting receive network buffer to 1049600 bytes
```

Solaris 10

```
setting receive network buffer to 65536 bytes
receive network buffer is 64240 bytes
```

Windows

```
setting receive network buffer to 1049600 bytes
setting receive network buffer to 1049600 bytes
```

The default value for this parameter is derived from the NetBackup data buffer size using the following formula:

For backup jobs: (<*data_buffer_size*> * 4) + 1024

For restore jobs: (<*data_buffer_size*> * 2) + 1024

For tape:

If the default value for the NetBackup data buffer size is 65536 bytes, the formula results in the following: a default NetBackup network buffer size of 263168 bytes for backups and 132096 bytes for restores.

For disk:

If the default value for the NetBackup data buffer size is 262144 bytes, the formula results in the following: a default NetBackup network buffer size of 1049600 bytes for backups and 525312 bytes for restores.

To set the network buffer size

1 Create the following files:

Linux/UNIX

```
/usr/openv/netbackup/NET_BUFFER_SZ
/usr/openv/netbackup/NET_BUFFER_SZ_REST
```

Windows

```
install_path\NetBackup\NET_BUFFER_SZ
install_path\NetBackup\NET_BUFFER_SZ_REST
```

2 Note the following about the buffer files:

These files contain a single integer that specifies the network buffer size in bytes. The value in each file must be within the value range that the operating system allows. Otherwise, the established connection may behave erratically. For example, to use a network buffer size of 64 kilobytes, the file would contain 65536. If the files contain the integer 0 (zero), the default operating system value for the network buffer size is used. More information about the buffer files is available:

https://www.veritas.com/content/support/en_US/article.100016112

If the `NET_BUFFER_SZ` file exists and the `NET_BUFFER_SZ_REST` file does not exist, `NET_BUFFER_SZ` specifies the network buffer size for backup and restores.

If the `NET_BUFFER_SZ_REST` file exists, its contents specify the network buffer size for restores.

If both files exist, the `NET_BUFFER_SZ` file specifies the network buffer size for backups. The `NET_BUFFER_SZ_REST` file specifies the network buffer size for restores.

Because local backup or restore jobs on the media server do not send data over the network, this parameter has no effect on those operations. This parameter is used only by the NetBackup media server processes that read from or write to the network, specifically, the `bptm` or `bpdm` processes. No other NetBackup process uses this parameter.

Network buffer size in relation to other parameters

The network buffer size parameter is the counterpart on the media server to the communications buffer size parameter on the client. The network buffer sizes need not be the same on all of your NetBackup systems for NetBackup to function properly. However, if the media server's network buffer size is the same as the client's communications buffer size, network throughput may improve.

Similarly, the network buffer size does not have a direct relationship to the NetBackup data buffer size.

See "About shared memory (number and size of data buffers)" on page 177.

The two buffers are separately tunable parameters. However, setting the network buffer size to a substantially larger value than the data buffer has achieved the best performance in many NetBackup installations.

Setting the NetBackup client communications buffer size

The NetBackup client has a tunable parameter to adjust the size of the network communications buffer. This buffer writes data to the network for backups.

This client parameter is the counterpart to the network buffer size parameter on the media server. The network buffer sizes are not required to be the same on all of your NetBackup systems for NetBackup to function properly. However, if the media server's network buffer size is the same as the client's communications buffer size, you may achieve better performance.

To set the communications buffer size parameter on Linux/UNIX clients

- Create the `/usr/openv/netbackup/NET_BUFFER_SZ` file.

  As with the media server, the file should contain a single integer that specifies the communications buffer size. Generally, performance is better when the value in the `NET_BUFFER_SZ` file on the client matches the value in the `NET_BUFFER_SZ` file on the media server. The value in each file must be within the value range that operating system allows. Otherwise, the established connection may behave erratically. More information about the buffer files is available:

  Best practices for NET_BUFFER_SZ and Buffer_size

  The `NET_BUFFER_SZ_REST` file is not used on the client. The value in the `NET_BUFFER_SZ` file is used for both backups and restores.

To set the communications buffer size parameter on Windows clients

1. From **Host Properties** in the NetBackup Administration Console, do the following: expand **Clients** and open the **Client Properties > Windows Client > Client Settings** dialog for the client on which the parameter is to be changed.
2. Enter the new value in the **Communications buffer** field.

   This parameter is specified in number of kilobytes. The default value is 128. However, for the best performance, set the value to 0 to allow the operating system to choose the buffer size.

   Because local backup jobs on the media server do not send data over the network, this parameter has no effect on these local operations. Only the NetBackup `bpbkar32` process uses this parameter. It is not used by any other NetBackup processes on a primary server, media server, or client.
3. If you modify the NetBackup buffer settings, test the performance of restores with the new settings.

About the NOSHM file

Each time a backup runs, NetBackup checks for the existence of the NOSHM file. No services need to be stopped and started for it to take effect. You might use NOSHM, for example, when the NetBackup server hosts another application that uses a large amount of shared memory, such as Oracle.

NOSHM is also useful for testing: both as a workaround while solving a shared memory issue, and to verify that an issue is caused by shared memory.

Note: NOSHM only affects operations when the client host is the media server.

NOSHM forces a local backup to run as though it were a remote backup. A local backup is a backup of a client that has a directly-attached storage unit. An example is a client that happens to be a primary server or media server. A remote backup passes the data across a network connection from the client to a primary server's or media server's storage unit.

A local backup normally has one or more client processes, for example `bpbkar`, that read from the disk and write into shared memory. A local backup also has a `bptm` process that reads from shared memory and writes to the storage media. A remote backup has one or more `bptm` (child) processes that read from a socket connection to bpbkar and write into shared memory. A remote backup also has a `bptm` (parent) process that reads from shared memory and writes to the storage media. NOSHM forces the remote backup model even when the client and the media server are the same system.

For a local backup without NOSHM, shared memory is used between `bptm` and `bpbkar`. Whether the backup is remote or local, and whether NOSHM exists or not, shared memory is always used between `bptm` (parent) and `bptm` (child).

Note: NOSHM does not affect the shared memory that `bptm` uses to buffer the data that is written to tape or disk. `bptm` uses shared memory for any backup, local or otherwise.

Using socket communications (the NOSHM file)

When a primary server or media server backs itself up, NetBackup uses shared memory to speed up the backup. In this case, NetBackup uses shared memory rather than socket communications to transport the data between processes. However, it may not be possible or desirable to use shared memory during a backup. In that case, you can use socket communications rather than shared memory to interchange the backup data.

To use socket communications

- Touch the following file:

  Linux/UNIX

  `/usr/openv/netbackup/NOSHM`

  Windows

  `install_path``\NetBackup\NOSHM`

  To touch a file means to change the file's modification and access times. The file name should not contain any extension.

NetBackup server performance in the data transfer path

To improve NetBackup server performance, consider the following factors regarding the data transfer path:

- See "About shared memory (number and size of data buffers)" on page 177.
- See "Changing parent and child delay values for NetBackup" on page 188.
- See "About NetBackup wait and delay counters" on page 187.
- See "Effect of fragment size on NetBackup restores" on page 204.

- See "Other NetBackup restore performance issues" on page 208.

About shared memory (number and size of data buffers)

The NetBackup media server uses shared memory to buffer data between the network and the tape drive or disk drive. (Or it buffers data between the disk and tape if the NetBackup media server and client are the same system.) The number and size of these shared data buffers can be configured on the NetBackup media server.

The number and size of the tape and disk buffers may be changed so that NetBackup optimizes its use of shared memory. A different buffer size may result in better throughput for high-performance tape drives. These changes may also improve throughput for other types of drives.

Buffer settings are for media servers only and should not be used on a pure primary server or client.

Note: Restores use the same buffer size that was used to back up the images being restored.

Default number of shared data buffers

Table 8-3 shows the default number of shared data buffers for various NetBackup operations.

Table 8-3 Default number of shared data buffers

NetBackup operation	Number of shared data buffers Linux/UNIX	Number of shared data buffers Windows
Non-multiplexed backup	30	30
Multiplexed backup	12	12
Restore that uses non-multiplexed protocol	30	30
Restore that uses multiplexed protocol	12	12
Verify	30	30
Import	30	30
Duplicate	30	30

Table 8-3 Default number of shared data buffers *(continued)*

NetBackup operation	Number of shared data buffers Linux/UNIX	Number of shared data buffers Windows
NDMP backup	30	30

Default size of shared data buffers

The default size of shared data buffers for various NetBackup operations is shown in Table 8-4.

Table 8-4 Default size of shared data buffers

NetBackup operation	Size of shared data buffers Linux/UNIX	Size of shared data buffers Windows
Non-multiplexed backup	64K (tape), 256K (disk)	64K (tape), 256K (disk)
Multiplexed backup	64K (tape), 256K (disk)	64K (tape), 256K (disk)
Restore, verify, or import	same size as used for the backup	same size as used for the backup
Duplicate	read side: same size as used for the backup write side: 64K (tape), 256K (disk)	read side: same size as used for the backup write side: 64K (tape), 256K (disk)
NDMP backup	64K (tape), 256K (disk)	64K (tape), 256K (disk)

On Windows, a single tape I/O operation is performed for each shared data buffer. Therefore, this size must not exceed the maximum block size for the tape device or operating system. For Windows systems, the maximum block size is generally 64K, although in some cases customers use a larger value successfully. For this reason, the terms "tape block size" and "shared data buffer size" are synonymous in this context.

Amount of shared memory required by NetBackup

You can use this formula to calculate the amount of shared memory that NetBackup requires:

- For tape:

Shared memory required = (number_data_buffers * size_data_buffers) * number_tape_drives * max_multiplexing_setting

- For disk:
  Shared memory required = (number_data_buffers * size_data_buffers)

For example, assume that the number of shared data buffers is 16, and the size of the shared data buffers is 64 kilobytes. Also assume two tape drives, and a maximum multiplexing setting of four. Following the formula, NetBackup requires 8 MB of shared memory:

```
(16 * 65536) * 2 * 4 = 8 MB
```

Be careful when changing these settings.

See "Testing changes made to shared memory" on page 187.

How to change the number of shared data buffers

You can change the number of shared data buffers by creating the following file(s) on the media server. In the files, enter an integer number of shared data buffers.

See "Notes on number data buffers files" on page 180.

- Linux/UNIX
  For tape:

  ```
  /usr/openv/netbackup/db/config/NUMBER_DATA_BUFFERS
  /usr/openv/netbackup/db/config/NUMBER_DATA_BUFFERS_RESTORE
  ```

  The value specified by `NUMBER_DATA_BUFFERS` determines the number of shared memory buffers for all types of backups and restores if none of the following `NUMBER_DATA_BUFFERS_xxxx` files exists.
  For disk:

  ```
  /usr/openv/netbackup/db/config/NUMBER_DATA_BUFFERS_DISK
  ```

  For multiple copies (Inline Copy):

  ```
  /usr/openv/netbackup/db/config/NUMBER_DATA_BUFFERS_MULTCOPY
  ```

  For the FT media server:

  ```
  /usr/openv/netbackup/db/config/NUMBER_DATA_BUFFERS_FT
  ```

  ```
  /usr/openv/netbackup/db/config/NUMBER_DATA_BUFFERS_RESTORE_FT
  ```

Note: To change the FT media server buffer sizes, contact Veritas support first.

- Windows
  For tape:

  install_path\NetBackup\db\config\NUMBER_DATA_BUFFERS
  install_path\NetBackup\db\config\NUMBER_DATA_BUFFERS_RESTORE

  The value specified by NUMBER_DATA_BUFFERS determines the number of shared memory buffers for all types of backups and restores if none of the following NUMBER_DATA_BUFFERS_xxxx files exists.
  For disk:

  install_path\NetBackup\db\config\NUMBER_DATA_BUFFERS_DISK

  For multiple copies (Inline Copy):

  install_path\NetBackup\db\config\NUMBER_DATA_BUFFERS_MULTCOPY

  For the FT media server:
  Note: The FT media server is not yet supported on Windows.

See "Testing changes made to shared memory" on page 187.

Notes on number data buffers files

Note the following points:

- The various number data buffers files must contain a single integer that specifies the number of shared data buffers NetBackup uses.
- If the NUMBER_DATA_BUFFERS file exists, its contents determine the number of shared data buffers to be used for multiplexed and non-multiplexed backups.
- The following NUMBER_DATA_BUFFERS files allow buffer settings for particular types of backups:
  - NUMBER_DATA_BUFFERS_DISK
  - NUMBER_DATA_BUFFERS_MULTCOPY
  - NUMBER_DATA_BUFFERS_FT

  The values specified in these files override either the NetBackup default number or the value that is specified in NUMBER_DATA_BUFFERS.
  For example, NUMBER_DATA_BUFFERS_DISK allows for a different value when you back up to disk instead of tape. If NUMBER_DATA_BUFFERS exists but NUMBER_DATA_BUFFERS_DISK does not, NUMBER_DATA_BUFFERS applies to tape and disk backups. If both files exist, NUMBER_DATA_BUFFERS applies to tape backups and NUMBER_DATA_BUFFERS_DISK applies to disk backups. If only NUMBER_DATA_BUFFERS_DISK is present, it applies to disk backups only.

- The NUMBER_DATA_BUFFERS file also applies to remote NDMP backups, but does not apply to local NDMP backups or to NDMP three-way backups.
  See "Note on shared memory and NetBackup for NDMP" on page 184.
- The NUMBER_DATA_BUFFERS_RESTORE file is only used for restore from tape, not from disk. If the NUMBER_DATA_BUFFERS_RESTORE file exists, its contents determine the number of shared data buffers for multiplexed restores from tape.
- The NetBackup daemons do not have to be restarted for the new buffer values to be used. Each time a new job starts, bptm checks the configuration file and adjusts its behavior.
- For a recommendation for setting NUMBER_DATA_BUFFERS_FT, refer to the following topic:
  See "Recommended number of data buffers for SAN Client and FT media server" on page 186.

How to change the size of shared data buffers

In general, default settings should work for most of the installation. However, if the server has enough free memory, you can change the size of shared data buffers by creating the following file(s) on the media server. In the files, enter an integer size in bytes for the shared data buffer. The integer should be a multiple of 1024 (a multiple of 32 kilobytes is recommended).

See "Notes on size data buffer files" on page 182.

- Linux/UNIX
  For tape:

  ```
  /usr/openv/netbackup/db/config/SIZE_DATA_BUFFERS
  ```

  The SIZE_DATA_BUFFERS value determines the shared memory buffer size for all types of backups if none of the following SIZE_DATA_BUFFERS_*xxxx* files exist.
  For tape (NDMP storage units):

  ```
  /usr/openv/netbackup/db/config/SIZE_DATA_BUFFERS_NDMP
  ```

  For disk:

  ```
  /usr/openv/netbackup/db/config/SIZE_DATA_BUFFERS_DISK
  ```

  The SIZE_DATA_BUFFERS_DISK file also affects NDMP to disk backups.
  For multiple copies (Inline Copy):

  ```
  /usr/openv/netbackup/db/config/SIZE_DATA_BUFFERS_MULTCOPY
  ```

For the FT media server:

```
/usr/openv/netbackup/db/config/SIZE_DATA_BUFFERS_FT
```

- Windows
  For tape:

  ```
  install_path\NetBackup\db\config\SIZE_DATA_BUFFERS
  ```

  The SIZE_DATA_BUFFERS value determines the shared memory buffer size for all types of backups if none of the following SIZE_DATA_BUFFERS_*xxxx* files exist.
  For tape (NDMP storage units):

  ```
  install_path\NetBackup\db\config\SIZE_DATA_BUFFERS_NDMP
  ```

  For disk:

  ```
  install_path\NetBackup\db\config\SIZE_DATA_BUFFERS_DISK
  ```

  The SIZE_DATA_BUFFERS_DISK file also affects NDMP to disk backups.
  For multiple copies (Inline Copy):

  ```
  install_path\NetBackup\db\config\SIZE_DATA_BUFFERS_MULTCOPY
  ```

  For the FT media server:
  Note: The FT media server is not yet supported on Windows.

See "Testing changes made to shared memory" on page 187.

Notes on size data buffer files

Note the following points:

- The various size data buffers files contain a single integer that specifies the size of each shared data buffer in bytes. The integer should be a multiple of 1024 (a multiple of 32 kilobytes is recommended).
  See "Size values for shared data buffers" on page 183.
- If the SIZE_DATA_BUFFERS file exists, its contents determine the size of shared data buffers to be used for multiplexed and non-multiplexed backups.
- The other SIZE_DATA_BUFFERS files (SIZE_DATA_BUFFERS_DISK, SIZE_DATA_BUFFERS_MULTCOPY, SIZE_DATA_BUFFERS_FT) allow buffer settings for particular types of backups. The values specified in these files override either the NetBackup default size or the value that is specified in SIZE_DATA_BUFFERS.

For example, SIZE_DATA_BUFFERS_DISK allows for a different value when you back up to disk instead of tape. If SIZE_DATA_BUFFERS exists but SIZE_DATA_BUFFERS_DISK does not, SIZE_DATA_BUFFERS applies to all backups. If both files exist, SIZE_DATA_BUFFERS applies to tape backups and SIZE_DATA_BUFFERS_DISK applies to disk backups. If only SIZE_DATA_BUFFERS_DISK is present, it applies to disk backups only.

- The SIZE_DATA_BUFFERS_DISK file also affects NDMP to disk backups.
- Perform backup and restore testing if the shared data buffer size is changed. If NetBackup media servers are not running the same operating system, test restores on each media server that may be involved in a restore operation. If a Linux/UNIX media server writes a backup to tape with a shared data buffer of 256 kilobytes, a Windows media server may not be able to read that tape.

Warning: Test restore as well as backup operations, to avoid the potential for data loss.

See "Testing changes made to shared memory" on page 187.

Size values for shared data buffers

Table 8-5 lists appropriate values for the various SIZE_DATA_BUFFERS files. The integer represents the size of one tape or disk buffer in bytes. For example, to use a shared data buffer size of 64 kilobytes, the file would contain the integer 65536.

These values are multiples of 1024. If you enter a value that is not a multiple of 1024, NetBackup rounds it down to the nearest multiple of 1024. For example, if you enter a value of 262656, NetBackup uses the value of 262144.

The NetBackup daemons do not have to be restarted for the parameter values to be used. Each time a new job starts, bptm checks the configuration file and adjusts its behavior.

Analyze the buffer usage by checking the bptm debug log before and after you alter the size of buffer parameters. Note that the bptm log applies to both tape backups and disk backups.

Table 8-5 Byte values for SIZE_DATA_BUFFERS_xxxx files

Kilobytes per data buffer	SIZE_DATA_BUFFER value in bytes
32	32768
64	65536

Table 8-5 Byte values for SIZE_DATA_BUFFERS_xxxx files *(continued)*

Kilobytes per data buffer	SIZE_DATA_BUFFER value in bytes
96	98304
128	131072
160	163840
192	196608
224	229376
256	262144

Important: the data buffer size equals the tape I/O size. Therefore the `SIZE_DATA_BUFFERS` value must not exceed the maximum tape I/O size that the tape drive or operating system supports. This value is usually 256 kilobytes or 128 kilobytes. Check your operating system and hardware documentation for the maximum values. Take into consideration the total system resources and the entire network. The Maximum Transmission Unit (MTU) for the LAN network may also have to be changed. NetBackup expects the value for `NET_BUFFER_SZ` and `SIZE_DATA_BUFFERS` to be in bytes. For 32K, use 32768 (32 x 1024).

Note: Some Windows tape devices cannot write with block sizes higher than 65536 (64 kilobytes). Some Windows media servers cannot read backups on a Linux/UNIX media server with `SIZE_DATA_BUFFERS` set to more than 65536. The Windows media server would not be able to import or restore images from media that were written with `SIZE_DATA_BUFFERS` greater than 65536.

Note: The size of the shared data buffers for a restore is determined by the size of the shared data buffers in use at the time the backup was written. Restores do not use the `SIZE_DATA_BUFFERS` files.

Note on shared memory and NetBackup for NDMP

The following tables describe how NetBackup for NDMP uses shared memory.

Table 8-6 shows the effect of NUMBER_DATA_BUFFERS according to the type of NDMP backup.

Table 8-6 NetBackup for NDMP and number of data buffers

Type of NDMP backup	Use of shared memory
Local NDMP backup or three-way backup	NetBackup does not use shared memory (no data buffers). `NUMBER_DATA_BUFFERS` has no effect.
Remote NDMP backup	NetBackup uses shared memory. You can use `NUMBER_DATA_BUFFERS` to change the number of memory buffers.

Table 8-7 shows the effect of SIZE_DATA_BUFFERS_NDMP according to the type of NDMP backup.

Table 8-7 NetBackup for NDMP and size of data buffers

Type of NDMP backup	Use of shared memory
Local NDMP backup or three-way backup	NetBackup does not use shared memory (no data buffers). You can use `SIZE_DATA_BUFFERS_NDMP` to change the size of the records that are written to tape. Use `SIZE_DATA_BUFFERS_DISK` to change record size for NDMP disk backup.
Remote NDMP backup	NetBackup uses shared memory. You can use `SIZE_DATA_BUFFERS_NDMP` to change the size of the memory buffers and the size of the records that are written to tape. Use `SIZE_DATA_BUFFERS_DISK` to change buffer size and record size for NDMP disk backup.

The following is a brief description of NDMP three-way backup and remote NDMP backup:

- In an NDMP three-way backup, the backup is written to an NDMP storage unit on a different NAS filer.
- In remote NDMP backup, the backup is written to a NetBackup Media Manager-type storage device.

More information is available on these backup types.

See the *NetBackup for NDMP Administrator's Guide*.

Recommended shared memory settings

Note: See Table 8-3 on page 177. for the default number of shared data buffers for various NetBackup operations.

The `SIZE_DATA_BUFFERS` setting for backup to tape is typically increased to 256 KB and `NUMBER_DATA_BUFFERS` is increased to 16. To configure NetBackup to use 16 x 256-KB data buffers, specify 262144 (256 x 1024) in `SIZE_DATA_BUFFERS` and 16 in `NUMBER_DATA_BUFFERS`.

Note that an increase in the size and number of the data buffers uses up more shared memory, which is a limited system resource. The total amount of shared memory that is used for each tape drive is:

Shared memory = (number_data_buffers * size_data_buffers) * number_tape_drives * max_multiplexing_setting

For two tape drives, each with a multiplexing setting of 4 and with 16 buffers of 256KB, the total shared memory usage would be:

(16 * 262144) * 2 * 4 = 32768 KB (32 MB)

If large amounts of memory are to be allocated, the kernel may require additional tuning to provide enough shared memory for NetBackup.

Make changes carefully, monitoring for performance changes with each modification. For example, an increase in the tape buffer size can cause some backups to run slower. Also, there have been cases with restore issues. After any changes, be sure to include restores as part of your validation testing.

Recommended number of data buffers for SAN Client and FT media server

For SAN Client Fibre Transport, the effective total number of data buffers is approximately twice the number of buffers that is specified for non-multiplexed backups. The reason is that the specified number of buffers are present on both the SAN Client and on the FT media server.

Note: It usually does not improve performance to increase memory buffers to a number that is significantly more than the SAN Client Fibre Transport default (16). Such an increase usually causes the majority of the buffers on either the client or server side to be empty.

The default setting of 16 allows 40 concurrent connections. Setting the buffers to 10 allows the maximum number of 64 concurrent connections. See *About Linux concurrent FT connections* in the NetBackup SAN Client and Fibre Transport Guide for more information.

Testing changes made to shared memory

After making any changes, it is vitally important to verify that the following tests complete successfully.

To test changes made to shared memory

1. Run a backup.
2. Restore the data from the backup.
3. Restore data from a backup that was created before the changes to the `SIZE_DATA_BUFFERS_xxxx` and `NUMBER_DATA_BUFFERS_xxxx` files.
4. Before and after altering the size or number of data buffers, examine the buffer usage information in the bptm debug log file.

   The values in the log should match your buffer settings. The relevant bptm log entries are similar to the following:

   ```
   12:02:55 [28551] <2> io_init: using 65536 data buffer size
   12:02:55 [28551] <2> io_init: CINDEX 0, sched bytes for
   monitoring = 200
   12:02:55 [28551] <2> io_init: using 8 data buffers
   ```

   or

   ```
   15:26:01 [21544] <2> mpx_setup_restore_shm: using 12 data
   buffers, buffer size is 65536
   ```

   When you change these settings, take into consideration the total system resources and the entire network. The Maximum Transmission Unit (MTU) for the local area network (LAN) may also have to be changed.

About NetBackup wait and delay counters

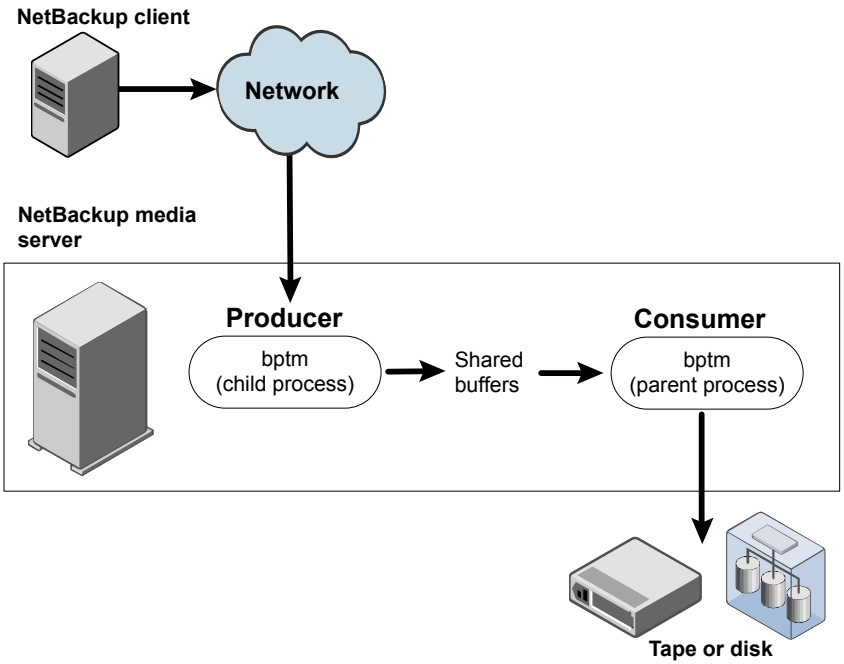
During a backup or restore operation, the NetBackup media server uses a set of shared data buffers to do the following: isolate the process of communicating with

the storage device (tape or disk) from the process of interacting with the client disk or network. Through the use of wait and delay counters, you can determine which process on the NetBackup media server has to wait more often: the data producer or the data consumer.

Achieving a good balance between the data producer and the data consumer processes is an important factor in achieving optimal performance from the NetBackup server component of the NetBackup data transfer path.

Figure 8-1 shows the producer-consumer relationship.

Figure 8-1 Producer-consumer relationship during a remote client backup



Changing parent and child delay values for NetBackup

You can modify the parent and child delay values for a process.

To change the parent and child delay values

1. Create the following files:

   Linux/UNIX

   ```
   /usr/openv/netbackup/db/config/PARENT_DELAY
   /usr/openv/netbackup/db/config/CHILD_DELAY
   ```

   Windows

   ```
   install_path\NetBackup\db\config\PARENT_DELAY
   install_path\NetBackup\db\config\CHILD_DELAY
   ```

   These files contain a single integer that specifies the value in milliseconds for the delay corresponding to the name of the file.

2. For example, for a parent delay of 50 milliseconds, enter 50 in the PARENT_DELAY file.

   See "About NetBackup wait and delay counters" on page 187.

About the communication between NetBackup client and media server

The communication process between the NetBackup client and the media server varies depending on the following:

- Whether the operation is a backup or restore
- Whether the operation involves a local client or a remote client

Table 8-8 NetBackup communication during backup and restore

Type of client (local or remote)	Communication process during backup and restore
Local client	When the NetBackup media server and the NetBackup client are part of the same system, the NetBackup client is referred to as a local client. ■ Backup of local client For a local client, the `bpbkar` (Linux/UNIX) or `bpbkar32` (Windows) process reads data from the disk during backup and places it in shared buffers. The `bptm` process reads the data from the shared buffer and writes it to tape or disk. ■ Restore of local client During a restore of a local client, the `bptm` process reads data from the tape or disk and places it in the shared buffers. The tar (Linux/UNIX) or tar32 (Windows) process reads the data from the shared buffers and writes it to disk. **Note:** Other processes may be used instead of bpbkar and bptm, depending on the data to be backed up or restored. See "Processes used in NetBackup client-server communication" on page 190.
Remote client	When the NetBackup media server and the NetBackup client are part of two different systems, the NetBackup client is referred to as a remote client. ■ Backup of remote client The `bpbkar` (Linux/UNIX) or `bpbkar32` (Windows) process on the remote client reads data from the disk and writes it to the network. Then a child `bptm` process on the media server receives data from the network and places it in the shared buffers. The parent `bptm` process on the media server reads the data from the shared buffers and writes it to tape or disk. ■ Restore of remote client During the restore of the remote client, the parent `bptm` process reads data from the tape or disk and places it into the shared buffers. The child `bptm` process reads the data from the shared buffers and writes it to the network. The tar (Linux/UNIX) or tar32 (Windows) process on the remote client receives the data from the network and writes it to disk. **Note:** Other processes may be used instead of bpbkar and bptm, depending on the data to be backed up or restored. See "Processes used in NetBackup client-server communication" on page 190.

Processes used in NetBackup client-server communication

This topic describes the logs that record the details about the NetBackup client-server communication. You can use these logs to adjust NetBackup client-server communication by means of wait and delay counters, as described in related topics.

Table 8-9 Logs used in NetBackup client-server communication

Log directory	Type of backup
Linux/UNIX log: /usr/openv/netbackup/logs/bpbkar Windows log: *install_path*\netbackup\logs\bpbkar	Standard, MS-Windows, FlashBackup, FlashBackup-Windows, Enterprise Vault, Lotus Notes, Domino, SharePoint, DB2 snapshot backups, Oracle RMAN PROXY/snapshot backups, Oracle Block Incremental Backups without RMAN, MS-SQL-Server snapshot backups
Linux/ UNIX log: /usr/openv/netbackup/logs/bpdb2 Windows log: *install_path*\netbackup\logs\bpdb2	DB2 stream-based backups. The `bpdb2` directory must be writable by the DB2 user (not only by the root user).
Linux/UNIX log: /usr/openv/netbackup/logs/dbclient Windows log: *install_path*\netbackup\logs\dbclient	Oracle RMAN stream-based backups. The `dbclient` directory must be writable by the Oracle user (not only by the root user).
Windows log: *install_path*\netbackup\logs\dbclient	MS-SQL-Server stream-based backups. (Not supported on UNIX or Linux.)
Linux/UNIX log: /usr/openv/netbackup/logs/exten_client Windows log: *install_path*\netbackup\logs\exten_client	DataStore/XBSA stream-based backups. The `exten_client` directory must be writable by the application that performs the backup operation or the restore operation.
Linux/UNIX log: /usr/openv/netbackup/logs/infxbsa	Informix stream-based backups. The `infxbsa` directory must be writable by the Informix user. (Not supported on Windows.)
Linux/UNIX log: /usr/openv/netbackup/logs/sybackup Windows log: *install_path*\netbackup\logs\sybackup	Sybase stream-based backups. The `sybackup` directory must be writable by the application that performs the backup operation or the restore operation.

Roles of processes during backup and restore

When a process attempts to use a shared data buffer, it first verifies that the next buffer is ready. A data producer needs an empty buffer, while a data consumer needs a full buffer.

The following table uses log directory names to represent the NetBackup processes.

Table 8-10 Roles of data producer and consumer

Operation	Data producer (Log directory name)	Data consumer (Log directory name)
Local backup	`bpbkar`, `bpdb2`, `dbclient`, `exten_client`, `infxbsa`, `sybackup`.	`bptm`
Remote backup	`bptm` (child)	`bptm` (parent)
Local restore	`bptm`	`tar` (Linux/UNIX) or `tar32` (Windows), `bpdb2`, `dbclient`, `exten_client`, `infxbsa`, `sybackup`.
Remote restore	`bptm` (parent)	`bptm` (child)

If the data consumer lacks a full buffer, it increments the wait and delay counters to indicate that it had to wait for a full buffer. After a delay, the data consumer checks again for a full buffer. If a full buffer is still not available, the data consumer increments the delay counter to indicate that it had to delay again. The data consumer repeats the delay and full buffer check steps until a full buffer is available.

This sequence is summarized in the following algorithm:

```
while (Buffer_IS_Not_Full) {
   ++Wait_Counter;
   while (Buffer_Is_Not_Full) {
      ++Delay_Counter;
      delay (DELAY_DURATION);
   }
}
```

If the data producer lacks an empty buffer, it increments the wait and delay counter to indicate that it had to wait for an empty buffer. After a delay, the data producer checks again for an empty buffer. If an empty buffer is still not available, the data producer increments the delay counter to indicate that it had to delay again. The

data producer repeats the delay and empty buffer check steps until an empty buffer is available.

The algorithm for a data producer has a similar structure:

```
while (Buffer_Is_Not_Empty) {
   ++Wait_Counter;
   while (Buffer_Is_Not_Empty) {
      ++Delay_Counter;
      delay (DELAY_DURATION);
   }
}
```

Analysis of the wait and delay counter values indicates whether the producer or the consumer process had to wait most often and for how long.

Four wait and delay counter relationships exist, as follows:

Table 8-11 Relationships between wait and delay counters

Relationship	Description
Data Producer >> Data Consumer	The data producer has substantially larger wait and delay counter values than the data consumer. The data consumer is unable to process the received data fast enough to keep the data producer busy. Investigate a means to improve the performance of the data consumer. For a backup, check if the data buffer size is appropriate for the tape or the disk drive being used. If the data consumer still has a substantially large value in this case, try increasing the number of shared data buffers to improve performance.
Data Producer = Data Consumer (large value)	The data producer and the data consumer have very similar wait and delay counter values, but those values are relatively large. This situation may indicate that the data producer and data consumer regularly attempt to use the same shared data buffer. Try increasing the number of shared data buffers to improve performance. See "Finding wait and delay counter values" on page 194.
Data Producer = Data Consumer (small value)	The data producer and the data consumer have very similar wait and delay counter values, but those values are relatively small. This situation indicates that there is a good balance between the data producer and data consumer. It should yield good performance from the NetBackup server component of the NetBackup data transfer path.

Table 8-11 Relationships between wait and delay counters *(continued)*

Relationship	Description
Data Producer << Data Consumer	The data producer has substantially smaller wait and delay counter values than the data consumer. The data producer is unable to deliver data fast enough to keep the data consumer busy. Investigate ways to improve the performance of the data producer. For a restore operation, check if the data buffer size is appropriate for the tape or the disk drive. If the data producer still has a relatively large value in this case, try increasing the number of shared data buffers to improve performance. See "How to change the number of shared data buffers" on page 179.

Of primary concern is the relationship and the size of the values. Information on determining substantial versus trivial values appears on the following pages. The relationship of these values only provides a starting point in the analysis. Additional investigative work may be needed to positively identify the cause of a bottleneck within the NetBackup data transfer path.

Finding wait and delay counter values

Wait and delay counter values can be found by creating debug log files on the NetBackup media server.

Note: The debug log files introduce additional overhead and have a small effect on the overall performance of NetBackup. This effect is more noticeable for a high verbose level setting. Normally, you should not need to run with debug logging enabled on a production system.

To find the wait and delay counter values for a local client backup

1. Activate debug logging by creating the appropriate log directories on the media server:

   Linux/UNIX

   For example:

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

   Windows

   ```
   install_path\NetBackup\logs\bpbkar
   install_path\NetBackup\logs\bptm
   ```

   The following topic lists log directories for processes that may be used in place of bpbkar, for database extensions:

   See "Processes used in NetBackup client-server communication" on page 190.

2. Execute your backup.

3. Consult the log for the data producer process.

   The line should be similar to the following, with a timestamp corresponding to the completion time of the backup:

   Example from the bpbkar log:

   ```
   ... waited 224 times for empty buffer, delayed 254 times
   ```

   In this example the wait counter value is 224 and the delay counter value is 254.

4. Look at the log for the data consumer process.

   The line should be similar to the following, with a timestamp corresponding to the completion time of the backup:

   ```
   ... waited for full buffer 1 times, delayed 22 times
   ```

   In this example, the wait counter value is 1 and the delay counter value is 22.

To find the wait and delay counter values for a remote client backup

1. Activate debug logging by creating this directory on the media server:

   Linux/UNIX

   ```
   /usr/openv/netbackup/logs/bptm
   ```

   Windows

   ```
   install_path\NetBackup\logs\bptm
   ```

2. Execute your backup.
3. Look at the log for the bptm process in:

   Linux/UNIX

   ```
   /usr/openv/netbackup/logs/bptm
   ```

   Windows

   ```
   install_path\NetBackup\logs\bptm
   ```

   Delays that are associated with the data producer (bptm child) appear as follows:

   ```
   ... waited for empty buffer 22 times, delayed 151 times, ...
   ```

   In this example, the wait counter value is 22 and the delay counter value is 151.

   Delays that are associated with the data consumer (bptm parent) appear as:

   ```
   ... waited for full buffer 12 times, delayed 69 times
   ```

   In this example the wait counter value is 12, and the delay counter value is 69.

To find the wait and delay counter values for a local client restore

1. Activate logging by creating the two directories on the NetBackup media server:

   Linux/UNIX

   ```
   /usr/openv/netbackup/logs/bptm
   /usr/openv/netbackup/logs/tar
   ```

   Windows

   ```
   install_path\NetBackup\logs\bptm
   install_path\NetBackup\logs\tar
   ```

   The following topic lists log directories for processes that may be used in place of tar, for database extensions:

   See "Processes used in NetBackup client-server communication" on page 190.

2. Execute your restore.

3. Look at the log for the data consumer (tar or tar32) in the tar log directory.

   The line should be similar to the following, with a timestamp corresponding to the completion time of the restore:

   ```
   ... waited for full buffer 27 times, delayed 79 times
   ```

   In this example, the wait counter value is 27, and the delay counter value is 79.

4. Look at the log for the data producer (`bptm`) in the bptm log directory.

   The line should be similar to the following, with a timestamp corresponding to the completion time of the restore:

   ```
   ... waited for empty buffer 1 times, delayed 68 times
   ```

   In this example, the wait counter value is 1 and the delay counter value is 68.

To find the wait and delay counter values for a remote client restore

1. Activate debug logging by creating the following directory on the media server:

   Linux/UNIX

   ```
   /usr/openv/netbackup/logs/bptm
   ```

   Windows

   ```
   install_path\NetBackup\logs\bptm
   ```

2. Execute your restore.
3. Look at the log for bptm in the bptm log directory.

   Delays that are associated with the data consumer (`bptm` child) appear as follows:

   ```
   ... waited for full buffer 36 times, delayed 139 times
   ```

   In this example, the wait counter value is 36 and the delay counter value is 139.

   Delays that are associated with the data producer (`bptm` parent) appear as follows:

   ```
   ... waited for empty buffer 95 times, delayed 513 times
   ```

   In this example the wait counter value is 95 and the delay counter value is 513.

Note on log file creation

When you run multiple tests, you can rename the current log file. Renaming the file causes NetBackup to create a new log file, which prevents you from erroneously reading the wrong set of values.

Deleting the debug log file does not stop NetBackup from generating the debug logs. You must delete the entire directory. For example, to stop bptm from logging, you must delete the bptm subdirectory. NetBackup automatically generates debug logs at the specified verbose setting whenever the directory is detected.

Use care when manipulating the bptm log files if the backup or restore is part of a multiplex (MPX) group that includes unrelated operations. In those instances, the bptm parent process opens the log file once at startup and receives a file descriptor from the operating system. The parent process and child processes write to that file descriptor until all current (and future) jobs that join the MPX group have completed. Unexpected consequences may result if the log file is renamed or deleted while the MPX group is still active.

If the log file is renamed, the file descriptor remains open against the renamed file. And if the next test job joins the same MPX group the new log entries appear in the renamed log file. If the log file is deleted, the file is no longer visible in the directory but the file descriptor remains open. If the next test job joins the same MPX group, the new log entries are written to the open file. Note that the user can no longer access the open file.

This behavior also applies to the MPX groups that have been running for multiple days. If the test job joins an MPX group that became active two days ago, the log entries are in the log from two days ago. If the bptm log directory did not exist two days ago, the bptm processes handling the backup do not generate any log entries.

If the bptm log directory did not exist two days ago, do one of the following:

- Wait for the MPX group to complete before starting the test job.
- Change either the storage unit, volume pool, or retention for the backup. The job is assigned to a drive and to media that is not already in use, and a new bptm parent process is started.

About tunable parameters reported in the bptm log

You can use the bptm debug log file to verify that the following tunable parameters have successfully been set to the desired values. You can use these parameters and the wait and delay counter values to analyze issues.

These additional values include the following:

Data buffer size	The size of each shared data buffer can be found on a line similar to: ... io_init: using 65536 data buffer size
Number of data buffers	The number of shared data buffers may be found on a line similar to: ... io_init: using 16 data buffers
Parent/child delay values	For the duration of the parent and child delays, the values in use can be found on a line similar to: ... io_init: child delay = 10, parent delay = 15 (milliseconds)

NetBackup media server network buffer size

The Network buffer size values on the media server appear in the debug log files in lines similar to the following. The first line indicates that the NET_BUFFER_SZ touch file exists. It also indicates that bptm has tried to change the network buffer space from the operating system default. (This first line may not be present.) The second line is always present and reflects the value that the operating system uses for the current backup or restore. If the two values differ, the operating system did not allow the change to be made. You should delete the NET_BUFFER_SZ file.

The bptm child process reads from the receive network buffer during a remote backup.

```
...setting receive network buffer to 263168 bytes
…receive network buffer is 49640 bytes
```

The bptm child process writes to the network buffer during a remote restore:

```
...setting send network buffer to 131072 bytes
…send network buffer is 131072 bytes
```

See "Setting the network buffer size for the NetBackup media server" on page 171.

Example of using wait and delay counter values

Suppose you wanted to analyze a local backup that has a 30-minute data transfer that is baselined at 5 MB per second. The backup involves a total data transfer of 9,000 MB. Because a local backup is involved, bptm is the data consumer. The data producer depends on the type of data that is backed up.

See "Processes used in NetBackup client-server communication" on page 190.

See "Roles of processes during backup and restore" on page 192.

Find the wait and delay values for the appropriate data producer process and for the consumer process (bptm) from the following:

See "Finding wait and delay counter values" on page 194.

For this example, suppose those values are the following:

Table 8-12 Examples for wait and delay

Process	Wait	Delay
bpbkar (Linux/UNIX) bpbkar32 (Windows)	29364	58033
bptm	95	105

These values reveal that `bpbkar` (or `bpbkar32`) is forced to wait by a `bptm` process that cannot move data out of the shared buffer fast enough.

Next, you can determine time lost due to delays by multiplying the delay counter value by the parent or child delay value, whichever applies.

In this example, the `bpbkar` (or `bpbkar32`) process uses the child delay value, while the `bptm` process uses the parent delay value. (The defaults for these values are 10 milliseconds for child delay and 15 milliseconds for parent delay.)

You can use the following equations to determine the amount of time lost due to these delays:

Table 8-13 Example delays

Process	Delay
bpbkar (Linux/UNIX) bpbkar32 (Windows)	58033 delays x 0.010 seconds = 580.33 seconds = 9 minutes 40 seconds
bptm	105 x 0.015 seconds = 1.6 seconds

Use these equations to determine if the delay for `bpbkar` (or `bpbkar32`) is significant. In this example, if this delay is removed, the resulting transfer time is:

30 minutes original transfer time - 9 minutes 40 seconds = 20 minutes 20 seconds (1220 seconds)

A transfer time of 1220 seconds results in the following throughput value:

9000 MB / 1220 seconds = 7.38 MB per second

7.38 MB per second is a significant increase over 5 MB per second. With this increase, you should investigate how the tape or disk performance can be improved.

You should interpret the number of delays within the context of how much data was moved. As the amount of moved data increases, the significance threshold for counter values increases as well.

Again, for a total of 9,000 MB of data being transferred, assume a 64-KB buffer.

You can determine the total number of buffers to be transferred using the following equation:

Number of kilobytes	9,000 x 1024 = 9,216,000 KB
Number of buffers	9,216,000 / 64 = 144,000

You can now express the wait counter value as a percentage of the total number of buffers:

bpbkar (Linux/UNIX), or bpbkar32 (Windows)	29364 / 144,000 = 20.39%
bptm	95 / 144,000 = 0.07%

In the 20 percent of cases where `bpbkar` (or `bpbkar32`) needed an empty shared data buffer, `bptm` has not yet emptied the shared data buffer. A value of this size indicates a serious issue. You should investigate as to why the data consumer (`bptm`) cannot keep up.

In contrast, the delays that `bptm` encounters are insignificant for the amount of data transferred.

You can also view the delay and wait counters as a ratio:

bpbkar (Linux/UNIX)	= 58033 delays / 29364 waits
bpbkar32 (Windows)	= 1.98

In this example, on average `bpbkar` (or `bpbkar32`) had to delay twice for each wait condition that was encountered. If this ratio is large, increase the parent or child delay to avoid checking for a shared data buffer in the correct state too often.

See “Changing parent and child delay values for NetBackup” on page 188.

Conversely, if this ratio is close to 1, reduce the applicable delay value to check more often, which may increase your data throughput performance. Keep in mind that the parent and child delay values are rarely changed in most NetBackup installations.

The preceding information explains how to determine if the values for wait and delay counters are substantial enough for concern.

Note: The wait and delay counters are related to the size of the data transfer. A value of 1,000 may be extreme when only 1 megabyte of data is moved. The same value may indicate a well-tuned system when gigabytes of data are moved. The final analysis must determine how these counters affect performance.

Issues uncovered by wait and delay counter values

You can correct issues by checking the following:

- `bptm-read` waits

  The `bptm` debug log contains messages such as the following:

  ```
  ...waited for full buffer 1681 times, delayed 12296 times
  ```

  The first number is the number of times that `bptm` waited for a full buffer: in other words, how many times the `bptm` write operations waited for data from the source. If the wait counter indicates a performance issue, a change in the number of buffers does not help. Multiplexing may help if the target of the backup is tape.

  See "Finding wait and delay counter values" on page 194.

  In general, if the full buffer delays are substantially higher than the empty buffer delays, the network communication between the backup clients and the media server should be examined first to make sure that the full buffer delay is not caused by slow ingestion by the client. If network communication is not the issue, then the next step should be to examine the job detail report using `bpdbjobs` command. If the job report shows long delays between backup job starts and I/O writing starts, then loading the last backup image may be taking too long. In that case, disabling the last image loading may be the answer.

- `bptm-write` waits

  The `bptm` debug log contains messages such as the following:

  ```
  ...waited for empty buffer 1883 times, delayed 14645 times
  ```

  The first number is the number of times that `bptm` waited for an empty buffer: the number of times `bptm` encountered data from the source faster than the data can be written to tape or disk. If the wait counter indicates a performance issue, reduce the multiplexing factor.

  See "Finding wait and delay counter values" on page 194.

  More buffers may help.

- `bptm` delays

  The `bptm` debug log contains messages such as the following:

  ```
  ...waited for empty buffer 1883 times, delayed 14645 times
  ```

The second number is the number of times that `bptm` waited for an available buffer. If the delay counter indicates a performance issue, investigate. Each delay interval is 30 microseconds.

Estimating the effect of multiple copies on backup performance

The **Multiple copies** option of the Schedule attributes takes one stream of data that the `bptm` buffers receive and writes to two or more destinations sequentially. (Previous releases of NetBackup referred to this option as Inline Copy, Inline Tape Copy, or ITC.) The time to write to multiple devices is the same as the time required to write to one device multiplied by the number of devices. The overall write speed, therefore, is the write speed of a single device divided by the number of devices.

The write speed of a backup device is usually faster than the read speed of the source data. Therefore, switching to multiple copies may not necessarily slow down the backup. The important figure is the write speed of the backup device: the native speed of the device multiplied by the compression ratio of the device hardware compression. For tape backups, this compression ratio can be approximated by looking at how much data is held on a single tape (as reported by NetBackup). Compare that amount of data with the uncompressed capacity of a cartridge.

For example:

LTO generations 6, 7, and 8 tape drives use a 2.5:1 compression ratio compared to the earlier LTO generations' 1.5:1 and 2:1 compression ratios. An LTO gen 6 drive has a native write speed of 160MB/s using the 2.5:1 compression ratio has a native write capacity of 160MB/s and a compression ratio of 400MB/s. Thus, 160MB/s * 2.5 = 400MB/s.

An LTO gen 6 cartridge has an uncompressed capacity of 2.5TB. The native write speed is 160MB/s. Using the 2.5:1 compression ratio, the compressed capacity will be 2.5TB * 2.5 = 6.25TB and a compressed write speed of 160MB/s * 2.5 = 400MB/s. Newer generations of LTO will follow this format.

If multiple copies to two LTO gen 6 drives are enabled, the overall write speed is 160/2 = 80MB/s.

If the backup normally runs at 80MB/s (the read speed of the source data is 80MB/s), multiple copies do not affect the backup speed. If the backup normally runs at 160MB/s, multiple copies reduce the speed of the backup.

Effect of fragment size on NetBackup restores

Fragment size can affect NetBackup restores for non-multiplexed and multiplexed images.

The fragment size affects where tape markers are placed and how many tape markers are used. (The default fragment size is 1 terabyte for tape storage units and 512 GB for disk.) As a rule, a larger fragment size results in faster backups, but may result in slower restores when recovering a small number of individual files.

The "Reduce fragment size to" setting on the Storage Unit dialog limits the largest fragment size of the image. By limiting the size of the fragment, the size of the largest read during restore is minimized, reducing restore time. The fragment size is especially important when restoring a small number of individual files rather than entire directories or file systems.

For many sites, a fragment size of approximately 10 GB results in good performance for backup and restore.

For a fragment size, consider the following:

- Larger fragment sizes usually favor backup performance, especially when backing up large amounts of data. Smaller fragments can slow down large backups. Each time a new fragment is created, the backup stream is interrupted.
- Larger fragment sizes do not hinder performance when restoring large amounts of data. But when restoring a few individual files, larger fragments may slow down the restore.
- Larger fragment sizes do not hinder performance when restoring from non-multiplexed backups. For multiplexed backups, larger fragments may slow down the restore. In multiplexed backups, blocks from several images can be mixed together within a single fragment. During restore, NetBackup positions to the nearest fragment and starts reading the data from there, until it comes to the desired file. Splitting multiplexed backups into smaller fragments can improve restore performance.
- During restores, newer, faster devices can handle large fragments well. Slower devices, especially if they do not use fast locate block positioning, restore individual files faster if fragment size is smaller. (In some cases, SCSI fast tape positioning can improve restore performance.)

Unless you have particular reasons for creating smaller fragments, larger fragment sizes are likely to yield better overall performance. For example, reasons for creating smaller fragments are the following: restoring a few individual files, restoring from multiplexed backups, or restoring from older equipment.

How fragment size affects restore of a non-multiplexed image

`bptm` positions to the media fragment and the actual tape block that contains the first file to be restored. If fast-locate is available, `bptm` uses that for the positioning.

If fast-locate is not available, `bptm` uses MTFSF/MTFSR (forward space filemark/forward space record) to do the positioning.

The first file is then restored.

After that, for every subsequent file to be restored, `bptm` determines where that file is, relative to the current position. It may be faster for `bptm` to position to that spot rather than to read all the data in between (if fast locate is available). In that case, `bptm` uses positioning to reach the next file instead of reading all the data in between.

If fast-locate is not available, `bptm` can read the data as quickly as it can position with MTFSR (forward space record).

Therefore, fragment sizes for non-multiplexed restores matter if fast-locate is NOT available. With smaller fragments, a restore reads less extraneous data. You can set the maximum fragment size for the storage unit on the Storage Unit dialog in the NetBackup Administration Console (**Reduce fragment size to**).

How fragment size affects restore of a multiplexed image on tape

`bptm` positions to the media fragment that contains the first file to be restored. If `fast_locate` is available, `bptm` uses that for the positioning. If `fast_locate` is not available, `bptm` uses MTFSF (forward space file mark) for the positioning. The restore cannot use "fine-tune" positioning to reach the block that contains the first file, because of the randomness of how multiplexed images are written. The restore starts to read, discarding data for other backup images included in the multiplexed group, and saving the data related to the image being restored. If the multiplex setting and number of co-mingled images were high at backup time, the restore may need to read and discard much more data than is actually restored.

The first file is then restored.

From that point, the logic is the same as for non-multiplexed restores, with one exception. If the current position and the next file position are in the same fragment, the restore cannot use positioning. It cannot use positioning for the same reason that it cannot use "fine-tune" positioning to get to the first file.

If the next file position is in a subsequent fragment (or on a different media), the restore uses positioning to reach that fragment. The restore does not read all the data in between.

Thus, smaller multiplexed fragments can be advantageous. The optimal fragment size depends on the site's data and situation. For multi-gigabyte images, it may be best to keep fragments to 1 gigabyte or less. The storage unit attribute that limits fragment size is based on the total amount of data in the fragment. It is not based on the total amount of data for any one client.

When multiplexed images are written, each time a client backup stream starts or ends, the result is a new fragment. A new fragment is also created when a checkpoint occurs for a backup that has checkpoint restart enabled. So not all fragments are of the maximum fragment size. End-of-media (EOM) also causes new fragment(s).

Some examples may help illustrate when smaller fragments do and do not help restores.

Example 1:

Assume you want to back up four streams to a multiplexed tape. Each stream is a single, 1-GB file. A default maximum fragment size of 1 TB has been specified. The resultant backup image logically looks like the following. 'TM' denotes a tape mark or file mark, which indicates the start of a fragment.

TM <4 gigabytes data> TM

To restore one of the 1-GB files, the restore positions to the TM. It then has to read all 4 GB to get the 1-GB file.

If you set the maximum fragment size to 1 GB:

TM <1 GB data> TM <1 GB data> TM <1 GB data> TM <1 GB data> TM

this size does not help: the restore still has to read all four fragments to pull out the 1 GB of the file being restored.

Example 2:

This example is the same as Example 1, but assume that four streams back up 1 GB of /home or C:\. With the maximum fragment size (**Reduce fragment size**) set to a default of 1 TB (assuming that all streams are relatively the same performance), you again end up with:

TM <4 GBs data> TM

Restoring the following

```
/home/file1
```

or

```
C:\file1
/home/file2
```

or

```
C:\file2
```

from one of the streams, NetBackup must read as much of the 4 GB as necessary to restore all the data. But, if you set **Reduce fragment size** to 1 GB, the image looks like the following:

TM <1 GB data> TM <1 GB data> TM <1 GB data> TM <1 GB data> TM

In this case, home/file1 or C:\file1 starts in the second fragment. `bptm` positions to the second fragment to start the restore of home/file1 or C:\file1. (1 GB of reading is saved so far.) After /home/file1 is done, if /home/file2 or C:\file2 is in the third or forth fragment, the restore can position to the beginning of that fragment before it starts reading.

These examples illustrate that whether fragmentation benefits a restore depends on the following: what the data is, what is being restored, and where in the image the data is. In Example 2, reducing the fragment size from 1 GB to half a GB (512 MB) increases the chance the restore can locate by skipping instead of reading, when restoring small amounts of an image.

Fragmentation and checkpoint restart

If the policy's Checkpoint Restart feature is enabled, NetBackup creates a new fragment at each checkpoint. It creates the fragment according to the **Take checkpoints every** setting. For more information on Checkpoint Restart, refer to the *NetBackup Administrator's Guide, Volume I*.

Other NetBackup restore performance issues

Table 8-14 Issues that affect NetBackup restore performance

Restore issues	Comments
NetBackup catalog performance	The disk subsystem where the NetBackup catalog resides has a large effect on the overall performance of NetBackup. To improve restore performance, configure this subsystem for fast reads. NetBackup binary catalog format provides scalable and fast catalog access. See "About the primary server NetBackup catalog" on page 34.
NUMBER_DATA_BUFFERS_RESTORE setting	This parameter can help keep other NetBackup processes busy while a multiplexed tape is positioned during a restore. An increase in this value causes NetBackup buffers to occupy more physical RAM. This parameter only applies to multiplexed restores. See "About shared memory (number and size of data buffers)" on page 177.
Index performance issues	Refer to *Indexing the Catalog for Faster Access to Backups* in the NetBackup Administrator's Guide, Volume I.
Multiplexing set too high	If multiplexing is too high, needless tape searching may occur. The ideal setting is the minimum needed to stream the drives.

Table 8-14 Issues that affect NetBackup restore performance *(continued)*

Restore issues	Comments
Restores from multiplexed database backups	NetBackup can run several restores at the same time from a single multiplexed tape, by means of the `MPX_RESTORE_DELAY` option. This option specifies how long in seconds the server waits for additional restore requests of files or raw partitions that are in a set of multiplexed images on the same tape. The restore requests received within this period are executed simultaneously. By default, the delay is 30 seconds. This option may be useful if multiple stripes from a large database backup are multiplexed together on the same tape. If the `MPX_RESTORE_DELAY` option is changed, you do not need to stop and restart the NetBackup processes for the change to take effect. When the request daemon on the primary server (`bprd`) receives the first stream of a multiplexed restore request, it triggers the `MPX_RESTORE_DELAY` timer. The timer starts counting the configured amount of time. `bprd` watches and waits for related multiplexed jobs from the same client before it starts the overall job. If another associated stream is received within the time-out period, it is added to the total job: the timer is reset to the `MPX_RESTORE_DELAY` period. When the time-out has been reached without `bprd` receiving an additional stream, the time-out window closes. All associated restore requests are sent to `bptm`. A tape is mounted. If any associated restore requests arrive, they are queued until the tape that is now "In Use" is returned to an idle state. If `MPX_RESTORE_DELAY` is not high enough, NetBackup may need to mount and read the tape multiple times to collect all header information for the restore. Ideally, NetBackup would read a multiplexed tape and collect all the required header information with a single pass of the tape. A single pass minimizes the restore time. See "Example of restore from multiplexed database backup (Oracle)" on page 209.

Example of restore from multiplexed database backup (Oracle)

Suppose that MPX_RESTORE_DELAY is not set in the bp.conf file, so its value is the default of 30 seconds. Suppose also that you initiate a restore from an Oracle RMAN backup that was backed up using 4 channels or 4 streams. You also use the same number of channels to restore.

RMAN passes NetBackup a specific data request, telling NetBackup what information it needs to start and complete the restore. The first request is received by NetBackup in 29 seconds, which causes the MPX_RESTORE_DELAY timer to be reset. The next request is received by NetBackup in 22 seconds; again the timer is reset. The third request is received 25 seconds later, resetting the timer a third time. But the fourth request is received 31 seconds after the third. Since the fourth request was not received within the restore delay interval, NetBackup starts three of the four restores. Instead of reading from the tape once, NetBackup queues the fourth restore request until the previous three requests are completed. Note that all of the

multiplexed images are on the same tape. NetBackup mounts, rewinds, and reads the entire tape again to collect the multiplexed images for the fourth restore request.

In addition to NetBackup's reading the tape twice, RMAN waits to receive all the necessary header information before it begins the restore.

If MPX_RESTORE_DELAY is longer than 30 seconds, NetBackup can receive all four restore requests within the restore delay windows. It collects all the necessary header information with one pass of the tape. Oracle can start the restore after this one tape pass, for better restore performance.

Set the MPX_RESTORE_DELAY with caution, because it can decrease performance if set too high. Suppose that the MPX_RESTORE_DELAY is set to 1800 seconds. When the final associated restore request arrives, NetBackup resets the request delay timer as it did with the previous requests. NetBackup must wait for the entire 1800-second interval before it can start the restore.

Therefore, try to set the value of MPX_RESTORE_DELAY so it is neither too high or too low.

NetBackup storage device performance in the data transfer path

This section looks at storage device functionality in the NetBackup data transfer path. Changes in these areas may improve NetBackup performance.

Tape drive wear and tear is much less, and efficiency is greater, if the data stream matches the tape drive capacity and is sustained. Most tape drives have slower throughput than disk drives. Match the number of drives and the throughput per drive to the speed of the SCSI/FC connection, and follow the hardware vendors' recommendations.

The following factors affect tape drives:

- Media positioning
  When a backup or restore is performed, the storage device must position the tape so that the data is over the read and write head. The positioning can take a significant amount of time. When you conduct performance analysis with media that contains multiple images, allow for the time lag that occurs before the data transfer starts.
- SCSI bus assignment
  Connect the tape drives to different SCSI buses. For example: If you have 8 tape drives, use a minimum of 4 SCSI cards and connect no more than 2 drives to each card.
- Tape streaming

If a tape device is used at its most efficient speed, it is "streaming" the data onto the tape. If a tape device is streaming, the media rarely has to stop and restart. Instead, the media constantly spins within the tape drive. If the tape device is not used at its most efficient speed, it may continually start and stop the media from spinning. This behavior is the opposite of tape streaming and usually results in a poor data throughput.

- Data compression
  Most tape devices support some form of data compression within the tape device itself. Compressible data (such as text files) yields a higher data throughput rate than non-compressible data, if the tape device supports hardware data compression.
  Tape devices typically come with two performance rates: maximum throughput and nominal throughput. Maximum throughput is based on how fast compressible data can be written to the tape drive when hardware compression is enabled in the drive. Nominal throughput refers to rates achievable with non-compressible data.

Note: NetBackup cannot set tape drive data compression. Follow the instructions that are provided with your OS and tape drive.

In general, tape drive data compression is preferable to client (software) compression. Client compression may be desirable for reducing the amount of data that is transmitted across the network for a remote client backup.
See "Compression and NetBackup performance" on page 219.

Chapter 9

Tuning other NetBackup components

This chapter includes the following topics:

When to use multiplexing and multiple data streams

For backup to tape, multiple data streams can reduce the time for large backups. The reduction is achieved by first splitting the data to be backed up into multiple streams. Then you use multiplexing, multiple drives, or a combination of the two for processing the streams concurrently. In addition, you can configure the backup so each physical device on the client is backed up by a separate data stream. Each data stream runs concurrently with streams from other devices, to reduce backup times.

Note: For best performance, use only one data stream to back up each physical device on the client. Multiple concurrent streams from a single physical device can adversely affect the time to back up the device: the drive heads must move back and forth between tracks that contain the files for the respective streams.

Multiplexing is not recommended for database backups when restore speed is of paramount interest or when your tape drives are slow.

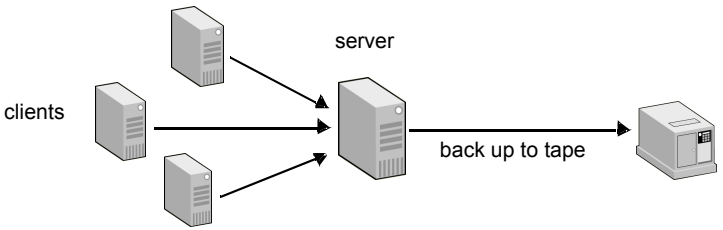
Backing up across a network, unless the network bandwidth is very broad, can nullify the ability to stream. Typically, a single client can send enough data to saturate a single 1Gb or 10Gb network connection. A gigabit network has the capacity to support network streaming for some clients. Multiple streams use more of the client's resources than a single stream. Veritas recommends testing to make sure of the following: that the client can handle the multiple data streams, and that the high rate of data transfer does not affect users.

Multiplexing and multiple data streams can be powerful tools to ensure that all tape drives are streaming. With NetBackup, both can be used at the same time. Be careful to distinguish between the two concepts, as follows.

Multiplexing writes multiple data streams to a single tape drive.

Figure 9-1 shows multiplexing.

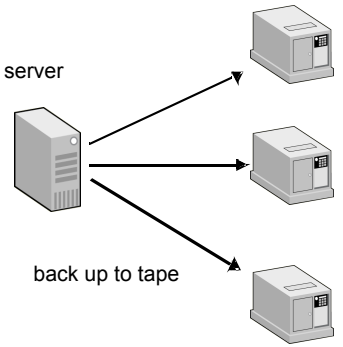
Figure 9-1 Multiplexing diagram



The multiple data streams feature writes multiple data streams, each to its own tape drive, unless multiplexing is used.

Figure 9-2 shows multiple data streams.

Figure 9-2 Multiple data streams diagram



Consider the following about multiplexing:

- Experiment with different multiplexing factors to find the one that is minimally sufficient for streaming.
  Find a setting at which the writes are enough to fill the maximum bandwidth of your drive: that setting is the optimal multiplexing factor. If you get 5 MB per second from each of the read streams, use a multiplexing factor of two to get the maximum throughput to an LTO 7 or LTO 8.
- Use a higher multiplexing factor for incremental backups.
- Use a lower multiplexing factor for local backups.
- Expect the duplication of a multiplexed tape to take longer if it is demultiplexed (unless "Preserve Multiplexing" is specified on the duplication). Without "Preserve Multiplexing," the duplication may take longer because multiple read passes of the source tape must be made. Using "Preserve Multiplexing," however, may affect the restore time (see next bullet).
- When you duplicate a multiplexed backup, demultiplex it.
  By demultiplexing the backups when they are duplicated, the time for recovery is significantly reduced.

Consider the following about multiple data streams:

- Do not use multiple data streams on single mount points.
  The multiple data streams feature takes advantage of the ability to stream data from several devices at the same time. Streaming from several devices permits backups to take advantage of Read Ahead on a spindle or set of spindles in

RAID environments. The use of multiple data streams from a single mount point encourages head thrashing and may result in degraded performance. Only conduct multistreamed backups against single mount points if they are mirrored (RAID 0). However, degraded performance is a likely result.

Effects of multiplexing and multistreaming on backup and restore

Note the following:

- Multiplexing
  To use multiplexing effectively, you must understand the implications of multiplexing on restore times. Multiplexing may decrease backup time for large numbers of clients over slow networks, but it does so at the cost of recovery time. Restores from multiplexed tapes must pass over all non-applicable data. This action increases restore times. When recovery is required, demultiplexing causes delays in the restore: NetBackup must search the tape to accomplish the restore.
  Restores should be tested to determine the impact of multiplexing on restore performance. Also, a smaller maximum fragment size when multiplexing may help restore performance.
  See "Effect of fragment size on NetBackup restores" on page 204.
  When you initially set up a new environment, keep the multiplexing factor low. A multiplexing factor of four or less does not highly affect the speed of restores, depending on the type of drive or system. If the backups do not finish within their assigned window, multiplexing can be increased to meet the window. However, a higher multiplexing factor provides diminishing returns as the number of multiplexing clients increases. The optimum multiplexing factor is the number of clients that are needed to keep the buffers full for a single tape drive.
  Set the multiplexing factor to four and do not multistream. Run benchmarks in this environment. Then you can begin to change the values until both the backup and restore window parameters are met.
- Multiple data streams
  The NEW_STREAM directive is useful for fine-tuning streams so that no disk subsystem is under-utilized or over-utilized.

How to improve NetBackup resource allocation

The following adjustments can be made to improve NetBackup resource allocation.

See "Improving the assignment of resources to NetBackup queued jobs" on page 216.

See "Sharing reservations in NetBackup" on page 216.

See "Disabling the sharing of NetBackup reservations" on page 216.

See "Disabling on-demand unloads" on page 218.

Improving the assignment of resources to NetBackup queued jobs

In certain situations, `nbrb` may take too long to process jobs that are waiting for drives. This delay may occur when many jobs are queued for resources and the jobs are completing faster than `nbrb` can re-use the released resources for new jobs.

To improve `nbrb` performance in this situation, see the `nbrbutil` command in the NetBackup Commands Reference Guide.

Also see the following article:

How to configure NetBackup Resource Broker (nbrb) to improve resource allocation performance

Sharing reservations in NetBackup

To improve performance, NetBackup allows shared reservations. NetBackup shares reservations by default. With shared reservations, multiple jobs can reserve the same media, though only one job can use it at a time. In other words, the second job does not have to wait for the first job to terminate. The second job can access the media as soon as the first job is done with it.

To enable the sharing of reservations

- Create the following file:

  On UNIX

  `/usr/openv/netbackup/db/config/RB_USE_SHARED_RESERVATIONS`

  On Windows

  `install_path``\Veritas\NetBackup\db\config\RB_USE_SHARED_RESERVATIONS`

Disabling the sharing of NetBackup reservations

In NetBackup, shared reservations are enabled by default.

See "Sharing reservations in NetBackup" on page 216.

In most cases, sharing reservations results in better performance.

However, it may be helpful to disable sharing reservations in the following case:

- Many duplication jobs are running (using a storage lifecycle policy, or Vault, or bpduplicate), and
- Many read media are shared between different duplication jobs

In this case, without shared reservations, one job runs and other jobs requiring the same media are queued because they cannot get a reservation. With shared reservations, the jobs can start simultaneously. However, with a limited set of resources (media/drive pair or disk drives), resources may bounce or "ping-pong" between different jobs as each job requests the resource.

For example, assume the following:

Two duplication jobs, job 1 and job 2, are duplicating backup images. Job 1 is duplicating images 1 through 5, and job 2 is duplicating images 6 through 9. The images are on the following media:

Table 9-1 Media required by jobs 1 and 2

Media used by job 1	Media used by job 2
Image 1 is on media A1	Image 6 is on media A2
Image 2 is on media A2	Image 7 is on media A2
Image 3 is on media A2	Image 8 is on media A2
Image 4 is on media A2	Image 9 is on media A3
Image 5 is on media A3	

In this example, both jobs require access to media A2. Without shared reservations, if job 1 gets the reservation first, job 2 cannot start, because it needs to reserve media A2. A2 is already reserved by job 1. With shared reservations, both jobs can start at the same time.

Assume, however, that only a few drives are available for writing. Also assume that job 1 begins first and starts duplicating image 1. Then job 2 starts using media A2 to duplicate image 6. Media A2 in effect bounces between the two jobs: sometimes it is used by job 1 and sometimes by job 2. As a result, the overall performance of both jobs may degrade.

You can use the following procedure to disable the sharing of reservations.

To disable the sharing of reservations

- Create the following file:

  On UNIX

  ```
  /usr/openv/netbackup/db/config/RB_DO_NOT_USE_SHARED_RESERVATIONS
  ```

  On Windows

  ```
  install_path\Veritas\NetBackup\db\config\RB_DO_NOT_USE_SHARED
  _RESERVATIONS
  ```

Disabling on-demand unloads

The NetBackup EMM service may ask the resource broker (nbrb) to unload drives even though the media unload delay has not expired. This request is called an on-demand unload. If allocating resources for a request is not possible without unloading the drive, EMM may ask nbrb to unload the drive.

It may be helpful to disable on-demand unloads when a series of small related backup jobs are scheduled (such as multiple NetBackup database agent jobs).

To disable on-demand unloads

- Create the following file:

  On UNIX

  ```
  /usr/openv/netbackup/db/config/RB_DISABLE_REAL_UNLOADS_ON_DEMAND
  ```

  On Windows

  ```
  install_path\Veritas\NetBackup\db\config\RB_DISABLE_REAL_UNLOADS_ON
  _DEMAND
  ```

Encryption and NetBackup performance

During the backup, encryption can be performed in any of the following ways, depending on your backup environment:

- The NetBackup client performs the encryption.
- The NetBackup media server performs the encryption.
- The tape drive performs the encryption, together with the NetBackup Key Management Service (KMS). The tape drive must have built-in encryption capability.

Table 9-2 describes the performance effect of each technology.

Table 9-2 Encryption options and NetBackup performance

Encryption option	Performance considerations
Client encryption (the **Encryption** option on the NetBackup policy attributes tab)	Data encryption (and compression) can be performed by the NetBackup client. (Use the encryption and compression options on the policy Attributes tab.) If the client has sufficient CPU resources to perform the encryption (plus the rest of its backup processing), client encryption can be an effective option. Note that when NetBackup client encryption is used, backups may run slower. How much slower depends on the throttle point in your backup path. If the network is the issue, encryption should not hinder performance. If the network is not the issue, then encryption may slow down the backup. If you multistream encrypted backups on a client with multiple CPUs, try to define one less stream than the number of CPUs. For example, if the client has four CPUs, define three or fewer streams for the backup. This approach can minimize CPU contention. See "Effect of encryption plus compression on NetBackup performance" on page 221. **Note:** Do not enable Encryption on the NetBackup policy attributes tab if backups are being written to a deduplication target, for example, an MSDP disk pool. Doing so will negatively impact the deduplication rate. Instead, enable MSDP encryption as described below.
Client encryption using MSDP	Backups that are being written to an MSDP disk pool can be encrypted using MSDP encryption. You have the option of encrypting a number of individual hosts or configuring encryption for all client direct clients. For additional information, see *Configuring encryption for MSDP backups* in the NetBackup Deduplication Guide.
Tape drive encryption, with the NetBackup Key Management Service (KMS)	Encryption that is performed by the tape drive has little or no effect on the backup performance. Use of this option requires the NetBackup Key Management Service (KMS). **Note:** The number of key groups in KMS is 100.

Compression and NetBackup performance

NetBackup supports two types of compression:

- Client compression (configured in the NetBackup policy)
- Tape drive compression (handled by the device hardware)

Consider the following when choosing the type of compression:

- If backup goes to MSDP, which does data compression by default, the client compression should not be enabled. Enabling the client compression may lead to poor deduplication rate in addition to potential performance impacts.
- The decision to use data compression should be based on the compressibility of the data itself.

  Note the following levels of compressibility, in descending order:

  - Plain text
    Usually the most compressible type of data.
  - Executable code
    May compress somewhat, but not as much as plain text.
  - Already compressed data
    Often, no further compression is possible.
  - Encrypted data
    May expand in size if compression is applied.
    See "Effect of encryption plus compression on NetBackup performance" on page 221.
- Tape drive compression is almost always preferable to client compression. Compression is CPU intensive, and tape drives have built-in hardware to perform compression.
- Avoid using both tape compression and client compression
  Compressing data that is already compressed can increase the amount of backed up data.
- Only in rare cases is it beneficial to use client (software) compression

  Those cases usually include the following characteristics:

  - The client data is highly compressible.
  - The client has abundant CPU resources.
  - You need to minimize the data that is sent across the network between the client and server.

  In other cases, however, NetBackup client compression should be turned off, and the hardware should handle the compression.
- Client compression reduces the amount of data that is sent over the network, but increases CPU usage on the client.
- On UNIX, the NetBackup client configuration setting MEGABYTES_OF_MEMORY may help client performance.
  This option sets the amount of memory available for compression.

Do not compress any files that are already compressed. If the data is compressed twice, refer to the NetBackup configuration option `COMPRESS_SUFFIX`. You can use this option to exclude files with certain suffixes from compression. Edit this setting through `bpsetconfig`.
See the *NetBackup Administrator's Guide, Volume II*.

How to enable NetBackup compression

Table 9-3 Tips on how to enable NetBackup compression

Type of compression	How to enable
Client compression	Select the compression option in the **NetBackup Policy Attributes** window.
Tape drive compression	Enabling tape drive compression depends on your operating system and the type of tape drive. Check with the operating system and drive vendors, or read their documentation to find out how to enable tape compression. Tips: With UNIX device addressing, these options are frequently part of the device name. A single tape drive has multiple names, each with a different functionality built into the name. (Multiple names are accomplished with major and minor device numbers.) If you address `/dev/rmt/2cbn` on Solaris, you get drive 2 hardware-compressed with no-rewind option. If you address `/dev/rmt/2n`, its function should be uncompressed with the no-rewind option. The choice of device names determines device behavior. If the media server is UNIX, there is no compression when the backup is to a disk storage unit. The compression options in this case are limited to client compression. If the media server with the disk storage unit is Windows, and the directory that is used by the disk storage unit is compressed, note: compression is used on the disk write as for any file writes to that directory by any application.

Effect of encryption plus compression on NetBackup performance

If a policy is enabled for both encryption and compression, the client first compresses the backup data and then encrypts it. When data is encrypted, it becomes

randomized, and is no longer compressible. Therefore, data compression must be performed before any data encryption.

Note the following information about encryption and compression:

- If the data provided to NetBackup is already encrypted or compressed, using NetBackup compression or encryption may be counterproductive and consume resources unnecessarily.
- If the backup data is stored with MSDP, which performs data compression by default, the client compression should not be enabled. Enabling the client compression may lead to poor deduplication rates in addition to potential poor general performance.

Information on NetBackup Java performance improvements

For performance improvements, refer to the following sections in the *NetBackup Administrator's Guide for UNIX and Linux, Volume I:*

- "Configuring the NetBackup-Java Administration Console,"
- "NetBackup-Java Performance Improvement Hints"

The *NetBackup Release Notes* may also contain information about NetBackup Java performance.

Information on NetBackup Vault

Information on tuning NetBackup Vault is available.

Refer to the "Best Practices" chapter of the *NetBackup Vault Administrator's Guide*.

Fast recovery with Bare Metal Restore

Veritas Bare Metal Restore (BMR) provides a simplified, automated method by which to recover an entire system (including the operating system and applications). BMR automates the restore process to ensure rapid, error-free recovery. This process requires one Bare Metal Restore command and then a system boot. BMR guarantees integrity and consistency and is supported for both UNIX and Windows systems.

Note: BMR requires the True image restore option. This option has implications for the size of the NetBackup catalog.

See "How to calculate the size of your NetBackup image database" on page 15.

How to improve performance when backing up many small files

NetBackup may take longer to back up many small files than a single large file.

Table 9-4 How to improve performance when backing up many small files

Try the following	Notes
Use the FlashBackup (or FlashBackup-Windows) policy type.	FlashBackup is a feature of NetBackup Snapshot Client. FlashBackup is described in the *NetBackup Snapshot Client Administrator's Guide*. See "How to improve FlashBackup performance" on page 224.
Windows: turn off virus scans.	Turning off scans may double performance.
Snap a mirror (such as with the FlashSnap method in Snapshot Client) and back that up as a raw partition.	Unlike FlashBackup, this type of backup does not allow individual file restore.
Turn off or reduce logging.	The NetBackup logging facility has the potential to affect the performance of backup and recovery processing. Logging is usually enabled temporarily, to troubleshoot a NetBackup problem. The amount of logging and its verbosity level can affect performance.
Make sure that the media server's network buffer size is the same as the client's communications buffer size	See "Setting the network buffer size for the NetBackup media server" on page 171. See "Setting the NetBackup client communications buffer size" on page 174.
Adjust the batch size for sending metadata to the catalog	See "Adjusting the batch size for sending metadata to the NetBackup catalog" on page 37.
Upgrade NIC drivers as new releases appear.	

Table 9-4 How to improve performance when backing up many small files *(continued)*

Try the following	Notes

Run a `bpbkar` throughput test	Run the following `bpbkar` throughput test on the client with Windows.
	`C:\Veritas\Netbackup\bin\bpbkar32 -nocont > NUL 2>`
	For example:
	`C:\Veritas\Netbackup\bin\bpbkar32 -nocont c:\ > NUL 2> temp.f`
	Run the following `bpbkar` throughput test on the client with UNIX.
	`/usr/openv/netbackup/bin/bpbkar -nocont -dt 0 -nofileinfo -nokeepalives` *`file system`* `> /dev/null`
	Where *file system* is the path being backed up.
	For example:
	`/usr/openv/netbackup/bin/bpbkar -nocont -dt 0 -nofileinfo -nokeepalives` *`file system`* `> /dev/null`

Optimize TCP/IP throughput	When initially configuring the Windows server, optimize TCP/IP throughput as opposed to shared file access.
Boost background performance on Windows versus foreground performance.	You can adjust processor scheduling by choosing how to allocate processor resources so that it is optimized to run programs (foreground services) or background services (for example, printing or backup) with more responsiveness.
Turn off NetBackup Client Job Tracker if the client is a system server.	See "NetBackup client performance in the data transfer path" on page 168.
Install appropriate patches	Regularly review the patch announcements for every server OS. Install patches that affect TCP/IP functions, such as correcting out-of-sequence delivery of packets.

How to improve FlashBackup performance

You can adjust NetBackup FlashBackup performance in the following ways.

Table 9-5 Tips for improving FlashBackup performance

Tips	Notes
Assign the snapshot cache device to a separate hard drive	If using the FlashBackup feature with a copy-on-write method such as nbu_snap, assign the snapshot cache device to a separate hard drive. A separate hard drive reduces disk contention and the potential for head thrashing. Refer to the *NetBackup Snapshot Client Administrator's Guide* for more information on FlashBackup configuration.
Adjust the FlashBackup read buffer	If the storage unit write speed is fast, reading the client disk may become a bottleneck during a FlashBackup raw partition backup. By default, FlashBackup (on UNIX) reads the raw partition using fixed 128 KB buffers for full backups and 32 KB buffers for incrementals. FlashBackup-Windows, by default, reads the raw partition using fixed 32 KB buffers for full backups and for incrementals. In most cases, the default read buffer size allows FlashBackup to stay ahead of the storage unit write speed. To minimize the number of I/O waits when reading client data, you can tune the FlashBackup read buffer size. Tuning this buffer allows NetBackup to read continuous device blocks up to 1 MB per I/O wait, depending on the disk driver. The read buffer size can be adjusted separately for full backup and for incremental backup. In general, a larger buffer yields faster raw partition backup (but see the following note). In the case of VxVM striped volumes, the read buffer can be configured as a multiple of the striping block size: data can be read in parallel from the disks, speeding up raw partition backup. **Note:** Resizing the read buffer for incremental backups can result in a faster backup in some cases, and a slower backup in others. Experimentation may be necessary to achieve the best setting. The result of the resizing depends on the following factors: ■ The location of the data to be read ■ The size of the data to be read relative to the size of the read buffer ■ The read characteristics of the storage device and the I/O stack. See "Adjusting the read buffer for FlashBackup and FlashBackup-Windows" on page 225.
Adjust the batch size for sending metadata to the catalog	See "Adjusting the batch size for sending metadata to the NetBackup catalog" on page 37.

Adjusting the read buffer for FlashBackup and FlashBackup-Windows

Use the following procedures to adjust the read buffer for NetBackup FlashBackup and FlashBackup-Windows raw partition backups.

To adjust the FlashBackup read buffer for UNIX and Linux clients

1. Create the following touch file on each client:

   ```
   /usr/openv/netbackup/FBU_READBLKS
   ```

2. Enter the values in the `FBU_READBLKS` file, as follows.

   On the first line of the file: enter an integer value for the read buffer size in blocks for full backups and/or for incremental backups. The defaults are 256 blocks (131072 bytes, or 128 KB) during full backups and 64 blocks (32768 bytes, or 32 KB) for incremental backups. The block size is equal to (KB size * 2), or (Number of bytes/512).

   To change both values, separate them with a space.

   For example:

   ```
   512 128
   ```

   This entry sets the full backup read buffer to 256 KB and the incremental read buffer to 64 KB.

   You can use the second line of the file to set the tape record write size, also in blocks. The default is the same size as the read buffer. The first entry on the second line sets the full backup write buffer size. The second value sets the incremental backup write buffer size. To set read buffer size and tape record write size to the same values, the file would read altogether as:

   ```
   512 128
   512 128
   ```

To adjust the FlashBackup-Windows read buffer for Windows clients

1. Click **Host Properties > Clients**, right-click on the client and select **Properties**.

   Click **Windows Client > Client Settings**.

2. For **Raw partition read buffer size**, specify the size of the read buffer.

   A read buffer size larger than the 32 KB default may increase backup speed. Results vary from one system to another; experimentation may be required. A setting of 1024 may be a good starting point.

   Note the following:

   - This setting applies to raw partition backups as well as to FlashBackup-Windows policies (including NetBackup for VMware).
   - This setting applies to full backups and to incremental backups.

Veritas NetBackup OpsCenter

For assistance in tuning Veritas NetBackup OpsCenter for better performance, refer to the following documents:

- NetBackup OpsCenter Administrator's Guide
- NetBackup OpsCenter Performance and Tuning Guide

Chapter 10

Tuning disk I/O performance

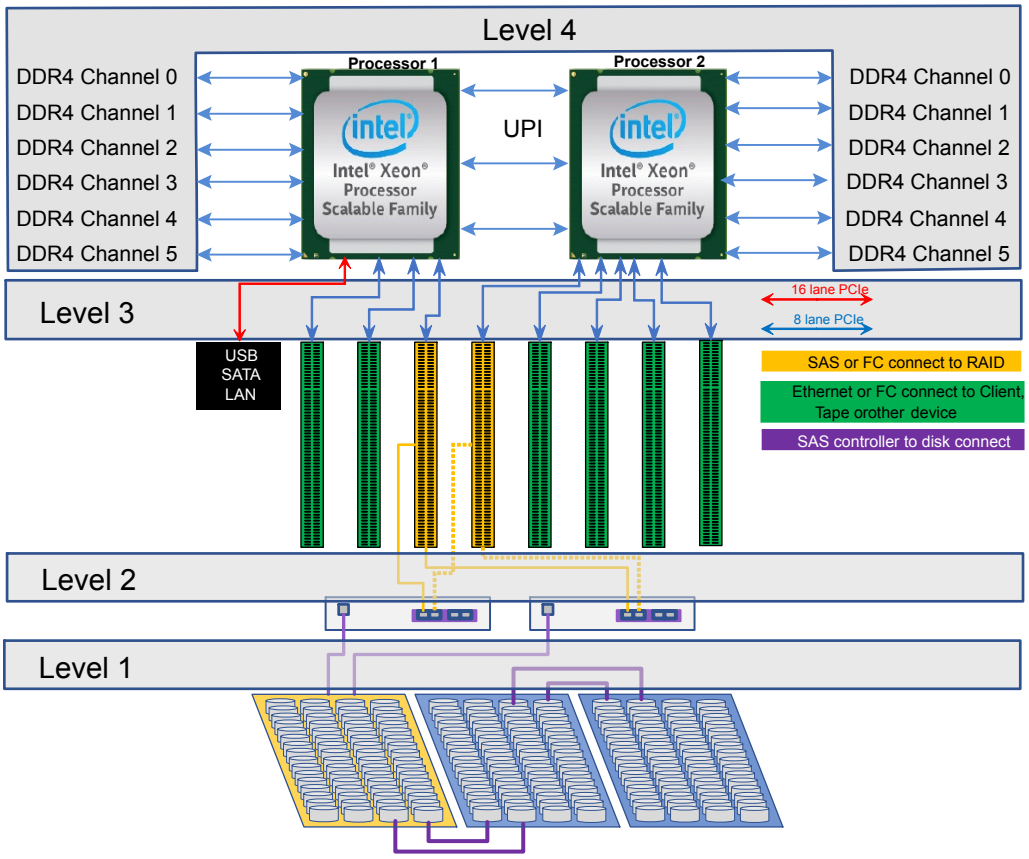
This chapter includes the following topics:

About NetBackup performance and the hardware hierarchy

The critical factors in NetBackup performance are not software-based. The critical factors are hardware selection and configuration. Hardware has roughly four times the weight that software has in determining performance.

Figure 10-1 shows the key hardware elements that affect performance, and the interconnections (levels) between them. The figure shows two disk arrays and a single non-disk device (tape, Ethernet connections, and so forth).

Figure 10-1 Performance hierarchy diagram

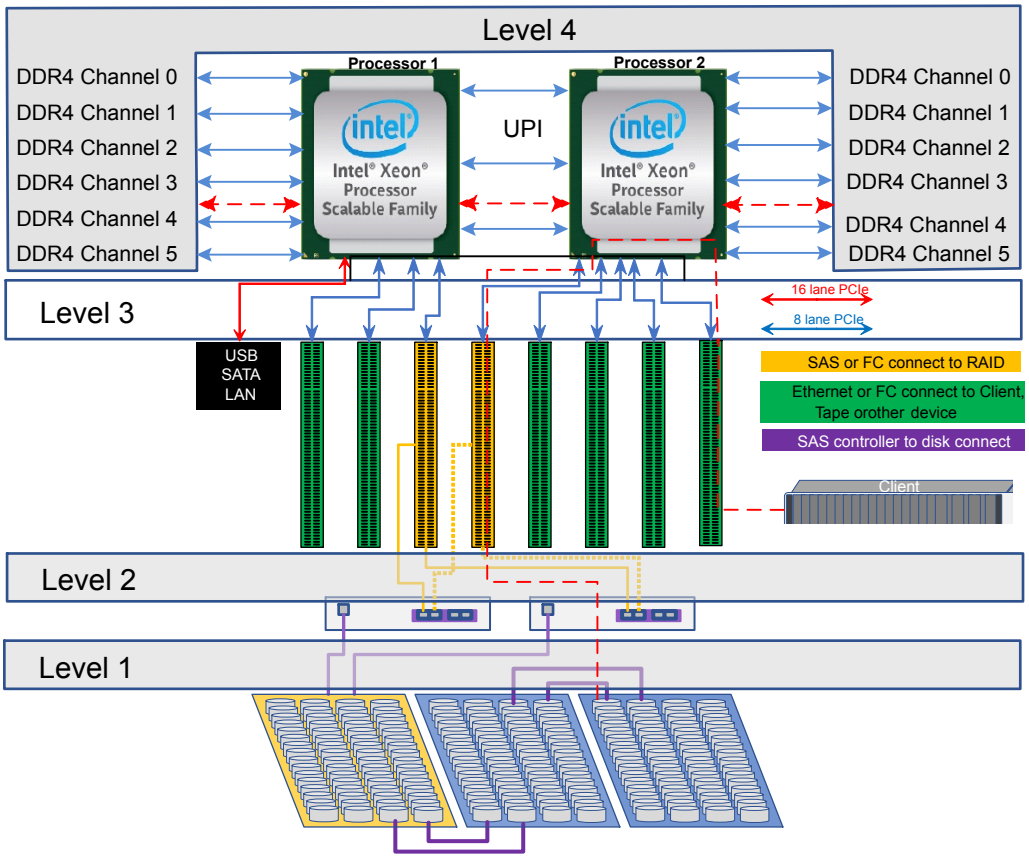


Performance hierarchy levels are described in later sections of this chapter.

In general, all data that goes to or comes from disk or SSD must pass through host memory.

Figure 10-2 includes a dashed line that shows the path that the data takes through a media server.

Figure 10-2 Data stream in NetBackup media server to arrays



The data moves up through an Ethernet NIC or Fibre Channel HBA on the client, to the Ethernet NIC or Fibre Channel HBA on the media server, acting as a target in this example, located in the rightmost PCIe slot. The data then moves directly into the Processor that is assigned to the PCIe slot and then to host memory. NetBackup then writes the new data to the appropriate location on the storage devices, via the PCIe NIC or HBA that interface with the RAID controllers. The data resides in the memory while NetBackup ascertains status: if it has been seen before or it is new and needs to be deduplicated. Efficiency of the PCIe lanes as components of the CPU drastically increases speed and lowers latency when compared to previous processors that required a companion chip to act as a bridge to the CPU.

About performance hierarchy level 1

Level 1 is the storage portion of a typical disk array. It can be populated with various size of Hard Disk Drives (HDD), typically 7200RPM SAS (Serial Attach SCSI (Small Computer System Interface)) that are dual ported to provide high availability. The storage can also be configured with Solid State Drives (SSD) for environments where access time is critical, such as concurrent read/write operations. As an example, backups with immediate replication to a remote site would benefit from SSDs. HDDs and SSDs are available in SAS configurations as SAS provides a redundant, reliable, mature protocol.

SAS runs at a speed of 12Gb/s in this example. The connectivity from the RAID controllers as well as the controllers themselves are noted with the purple lines and outlines. The connectivity from the controllers to the drives are 12Gb and the drives have a dual ported connection. Each of the drive SAS ports are connected via a SAS fabric with “expander” that allows for a high number of concurrent connections to disks, SSDs, tapes, and RAID controllers. The dual port connection allows for dual controller operation, the whole of which produces a highly reliable solution with 99.999% (5 nines) of availability. This solution is used primarily on the higher capacity systems that require speed and availability as key requirements.

When using disk drives of 1TB or higher capacity, it is highly recommended that users build their systems with RAID 6/dual parity or RAID 10 to ensure that no data is lost due to disk failures. Systems with the size of drives equal to or greater than 1TB take a relatively long time to rebuild to a hot spare that should be included in every system. The long rebuild time creates a need to plan against a second drive failure during rebuild. RAID 6 addresses this as its configuration provides two parity drives.

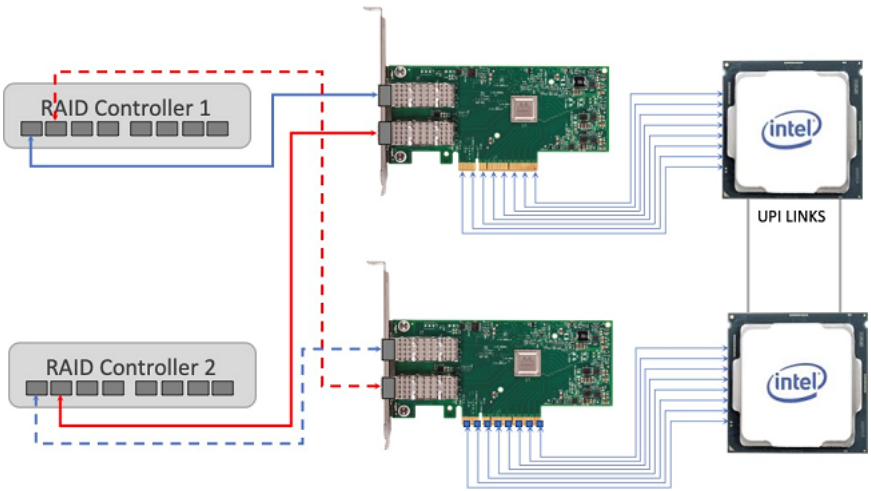
When creating storage, note that NetBackup supports volumes of up to 32 TiB and a maximum of 6 volumes per NetBackup Media server, so decide on your storage solution with this in mind. As an example, if 4TB drives were used and the volume had 11 drives configured in a RAID6 Logical Unit Number (LUN) the effective capacity in TB would be 4 * (11-2) = 36TB. In TebiBytes (TiB) it converts to 32.74 TiB.

This same type of solution can be created with SSDs. A major consideration is that the speed of rebuild for a disk drive is much longer than an SSD. Rebuilds with a 7.68TB SSD are approximately, 2.25 hours as compared to the 4TB disk drive noted earlier which, if there are numerous volumes being managed by dual controllers, can stretch to 40+ hours. Because of this difference in rebuild speed, with SSD drives, it is possible to build RAID5 volumes and have a fast enough rebuild to remove the need for dual parity. Use cases for SSD include re-hydration of deduplicated data to save to tape and operations, such as multiple concurrent backups and replications.

About performance hierarchy level 2

Level 2 contains the connectivity options to enable communication with the external storage and clients. This level includes Ethernet Network Interface Cards (NIC), Fibre Channel (FC) Host Bus Adapters (HBA) Serial Attach SCSI (SAS) RAID controllers, and SAS HBA. Typical Server level platforms utilize 8 lane PCIe slots for connecting HBAs and NICs. Some systems will have 2 each 16 lane slot PCIe that double the performance of the individual slot but this limits the total number of slots available for the peripherals. Experience has shown that performance with 8 lane slots, either PCIe3 at 7.877GB/s or PCIe4 at 15.754GB/s is best from a cost and performance perspective. The best solution for the storage attached with Redundant Controllers and dual processor systems is to route two SAS or Fibre Channel HBAs from each of the processors in a dual CPU compute node. The complimentary best practice is to route the Fibre channel or SAS from the two ports on each Host Bus Adapter to each of the controllers. See below for a diagram of the attach.

Figure 10-3 Performance hierarchy diagram



About performance hierarchy level 3

Because of the number and the speed of the PCIe lanes, as noted in Figure 8-2, the ability of a single system to address a large number of clients and provide large capacity storage, is made possible. In 2021, the Intel Architecture will migrate to the Generation 4 of the PCIe protocol doubling the transfer rate on PCIe lanes from 0.985GB/s to 1.969 GB/s. Current generation of systems are largely 8 lane Generation 3 with a bandwidth of 7.877 GB/s. This speed allows for 2 ports of up

to 25Gb Ethernet and up to 2 ports of 32Gb Fibre Channel HBAs to operate on the Generation 3 based systems.

Generation 4 PCIe includes the doubling of speed with an increase in the number of PCIe lanes per processor. The previous generation of Intel Xeon processors had 40 lane and the new Gen 4 processors have 64 lanes. AMD processors with the advent of the “Rome” have 128 Gen 4 lanes. When used as a dual processor system the interconnect between Intel processors is done with 2 or more “Ultra Path Interconnect” (UPI) links and the AMD Rome processors use 64 PCIe Gen 4 lanes from each processor as the interconnect. In both dual processor solutions, the number of available PCIe lanes is 128. If we compare the speed and number of lanes to Gen 3, there is a 3.2x performance improvement for the PCIe on the dual Intel processor architecture.

This increment creates a new use case for the Ethernet and Fibre Channel population. With an 8 lane PCIe slot we can increment to 4 each 25Gb ports per Ethernet NIC and 4 each 32Gb Fiber Channel HBA. Coupled with the larger number of cores per processor, the number of concurrent backups will increase on a per Media Server basis. This capability that the use of higher count ports per NIC or HBA allows the user to halve the number of PCIe cards on the system. This reduces cost and allows for a higher airflow that will maximize the life of the components. PCIe Gen5 is in process and will likely be the standard in the calendar year 2024. This will allow a doubling of performance and it is expected that quad port 100Gb Ethernet NICs will be the norm.

About performance hierarchy level 4

This section considers the population of the Dynamic Random Access Memory (DRAM) and the processor, specifically the number of cores should be populated in the system. Many people look to speed of the processor as the primary determinate of performance. However, backup and restore, particularly multiple operations of each are more reliant on the number of cores as each core can be performing a single or small number of operations.

DRAM population has a very close relationship to the amount of MSDP data to be addressed. The easy way to calculate how much RAM is required in the system is to allocate 1 Gigabyte of RAM for every 1 Tebibyte of MSDP data in the system. As an example, if creating a system that will have the non-appliance NetBackup maximum of 256TiB of MSDP storage, the system should have at minimum 256GB of RAM. The reason for this is the implementation of the SHA-2 HASH function as described in Chapter 3:

See “Fingerprint lookup for deduplication” on page 55.

If you are building a system with the 256 TiB of MSDP storage, it is a good idea to ramp up the RAM to 384 or 512GB to ensure no performance degradation when the storage starts to fill up with deduplicated data.

Lastly, the number of processor cores to allocate to the system is dependent on the type of backup jobs to be done. For instance, if concurrently backing up large number of VMware instances, then a larger number of cores is a good idea as there can be more concurrent jobs completed in a shorter time frame. On the other side, if backing up small number of large files the number of cores can be lower.

One method that is a good way to size your solution is, if building a higher performance system use the cost of approximately $1300.00 per processor. You can find this information at https://en.wikipedia.org/wiki/List_of_Intel_Xeon_processors. As an example, if we look at a current (as of the publishing of this document) list price there is a 24 core Xeon Gold 5318Y processor for $1283.00. Two of these would create a good solution with a total of 48 cores.

For mid-range systems, use the $700.00 mark. In this instance there is a Xeon Silver 4314 that has an MSRP of $694 and has 16 cores per processor. This solution would provide 32 cores, enough for the vast majority of systems, and able to accommodate tight budgets.

Summary of performance hierarchies

When you are putting together systems, remember that you are doing this to ensure that your data is available and has enough bandwidth/speed to meet the performance requirements. For this reason, invest in quality disk or SSD storage as criteria 1.

Next, configure enough bandwidth on your network to handle the peak amount of traffic allowing you to process the maximum concurrent jobs defined in the service level agreement.

Lastly, configure enough processor cores that are needed for the type of data to be backed up. Do not under populate the amount of DRAM allocated to the compute node. For example, we recommend at least 1GB of DRAM for every TB of storage for running deduplication.

More recommendations on the hardware design of the NetBackup compute node is available:

See “NetBackup hardware design and tuning considerations” on page 44.

Notes on performance hierarchies

The hardware components between interconnection levels can also affect bandwidth, as follows:

- A drive has sequential access bandwidth and average latency times for seek and rotational delays.

  Drives perform optimally when doing sequential I/O to disk. Non-sequential I/O forces movement of the disk head (that is, seek and rotational latency). This movement is a huge overhead compared to the amount of data transferred. The more non-sequential I/O done, the slower it becomes.

  Simultaneously reading or writing more than one stream results in a mix of short bursts of sequential I/O with seek and rotational latency in between. This situation significantly degrades overall throughput, especially when the drive approaches 85% capacity. Different drive types have different seek and rotational latency specifications. Therefore, the type of drive has a large effect on the amount of degradation.

  From best to worst, such disk drives are Serial Attach SCSI (SAS). “SATA” Serial ATA drives are rarely used for enterprise backup systems. SATA Solid-State Disks (SSDs) have access times at least twice that of their rotating disk counterparts and a 0.1 versus 12-millisecond access time. NVMe Gen4 SSDs have speeds of 3.8Gb/s writes and 7GB/s reads. For large repositories of data that is kept available, but if speed of retrieval is not a prime concern disk drive is the largest most cost efficient storage medium.

- A RAID controller has cache memory of varying sizes. The controller also does the parity calculations for RAID-5 or RAID-6. RAID-5 can be used on SSD systems as the rebuild time is fast enough to not have a concurrent failure. RAID-6 or RAID-10 is required for Disk-based storage.

- A PCIe card can be limited either by the speed of the ports or the clock rate to the PCIe slot. More information about the size and speed of the PCIe configurations is available:

  See Table 3-1 on page 45.
  See Figure 3-1 on page 47.

Memory can be a limit if there is intensive non-I/O activity in the system. Ensure that you follow the 1GB of memory to 1TB of MSDP storage so that the SHA-2 encryption is efficiently handled.

While CPU performance contributes to all performance, it is not the bottleneck in most modern systems for I/O intensive workloads. Very little work is done at that level. The CPU must execute a read operation and a write operation, but those operations do not take up much bandwidth. An exception fingerprint calculation in MSDP deduplication is CPU intensive. High number of concurrent streams with a high deduplication ratio can create a CPU bottleneck. Plan for at least one CPU thread per backup stream on lower powered systems. Up to four concurrent streams per core (two threads per core) on Intel processors are possible with small and numerous file backups

Hardware examples for better NetBackup performance

These examples are not intended as recommendations for your site. The examples illustrate various hardware factors that can affect NetBackup performance.

Example 1

A general hardware configuration can have dual 16-gigabit Fibre Channel ports on a single PCI card.

In such a case, the following is true:

- Potential bandwidth is approximately 14 Gb per second.
- For maximum performance, the card must be plugged into at least an 8-lane 3.0 PCIe slot.

Example 2

The next example shows a pyramid of bandwidth potentials with aggregation capabilities at some points.

Suppose you have the following hardware:

- 1x quad 1-gigabit ethernet
- 6x 10Gb Ethernet ports
- 4x 16Gb Fibre Channel (2 HBAs)
- 2x4x12Gb SAS port 12 bay JBOD with 8 TB 7200 RPM SAS-3 Disk drives. Up to 4 JBODs can be used to aggregate storage up to 256TiB with 8TB drive population.
- 1x Dual Socket Intel Scalable Server with 6 or more PCIe Slots, 64, 256 or 512 GB DDR4 ECC DRAM, two 16 core processors, Internal SAS-3 RAID controller with 2 internal and 2 external 4x12Gb ports

In this case, the following is one way to assemble the hardware so that no constraints limit throughput:

- The quad 1-gigabit ethernet card can do approximately 400 MB per second throughput.
- Each 10Gb Ethernet port can do approximately 1 GB per second for a total of 6 GB aggregated throughput.
- Each 16Gb Fibre Channel port can do approximately 1.75 GB per second for a total of 7 GB per second.

- The RAID controller / Disk JBOD combination can achieve 1.9GB per second with sequential data on a single JBOD.

Note the following:

- Each card can therefore move 2 to 3.5GB per second. With five cards, the total potential is 13GB per second.
- It should be noted that this level of performance would not be reached in a typical backup environment.
- External influences such as network traffic, will have an effect on overall performance.
- Utilizing the configuration detailed above will provide a high level of performance, especially if at least the "1GB RAM to 1TB MSDP" data rule is followed closely.