



# NetBackup Disk Based Data Protection Options

*Alex Davies*

## Contents

<b>1.0</b>	<b>INTRODUCTION</b>	<b>4</b>
1.1	GLOSSARY	5
<b>2.0</b>	<b>NETBACKUP ENTERPRISE DISK FOUNDATION</b>	<b>6</b>
2.1	MEDIA LOAD SERVER BALANCING AND FAILOVER	6
2.2	INTELLIGENT DISK CAPACITY MANAGEMENT	7
2.3	STORAGE LIFECYCLE POLICIES (LIFECYCLES)	7
2.3.1	<i>Capacity Managed Retention and Staging</i>	8
2.3.2	<i>Data Classifications</i>	8
2.4	DISK POOLS	9
2.5	INTELLIGENT DISK STORAGE DEVICES	11
<b>3.0</b>	<b>NETBACKUP DISK TYPES</b>	<b>13</b>
3.1	BASICDISK	13
3.1.1	<i>What is BasicDisk?</i>	13
3.1.2	<i>What's good about BasicDisk?</i>	13
3.1.3	<i>What are the downsides of BasicDisk?</i>	13
3.1.4	<i>When should I consider using BasicDisk?</i>	13
3.2	ADVANCEDDISK	14
3.2.1	<i>What is AdvancedDisk?</i>	14
3.2.2	<i>What are the benefits of AdvancedDisk?</i>	14
3.2.3	<i>What are the downsides of AdvancedDisk?</i>	14
3.2.4	<i>When should I consider using AdvancedDisk?</i>	14
3.3	SHAREDISK	15
3.3.1	<i>What is SharedDisk?</i>	15
3.3.2	<i>What's good about SharedDisk?</i>	15
3.3.3	<i>What's the downside of SharedDisk?</i>	15
3.3.4	<i>When should I consider using SharedDisk?</i>	15
3.4	VIRTUAL TAPE LIBRARY (VTL)	16
3.4.1	<i>What is a VTL?</i>	16
3.4.2	<i>What's good about a VTL?</i>	16
3.4.3	<i>What are the downsides of a VTL?</i>	16
3.4.4	<i>When should I consider using a VTL?</i>	17
3.5	OPENSTORAGE	17
3.5.1	<i>What is OpenStorage?</i>	17
3.5.2	<i>What's good about OpenStorage?</i>	17
3.5.3	<i>What are the downsides of OpenStorage?</i>	18
3.5.4	<i>When should I consider using OpenStorage?</i>	18
3.6	PUREDISK DEDUPLICATION OPTION	18
3.6.1	<i>What is PureDisk Deduplication Option?</i>	18
3.6.2	<i>What's good about PureDisk Deduplication Option?</i>	18
3.6.3	<i>What are the downsides of PureDisk Deduplication Option?</i>	19
3.6.4	<i>When should I consider using PureDisk Deduplication Option?</i>	19
<b>4.0</b>	<b>IMPLEMENTING A DISK BASED SOLUTION</b>	<b>20</b>
4.1	CHOOSING THE RIGHT DISK TYPE FOR THE JOB	20
4.1.1	<i>Capability</i>	20
4.1.2	<i>Cost of ownership</i>	21
4.1.3	<i>Compatibility</i>	22
4.2	DISK BASED SOLUTION EXAMPLES	22
4.2.1	<i>Balanced backup for optimal throughput</i>	22
4.2.2	<i>Combined Main and Satellite Site Protection Using PureDisk</i>	23
4.2.3	<i>Highly Available Dual Site/Single Domain Configuration</i>	24



## 1.0 Introduction

The move towards disk based data protection solutions is the single greatest market trend in data protection today. Disk based data protection solutions improve data protection performance, reliability, and disaster recovery capabilities. They lower cost, decrease backup windows, increase flexibility, and provide unique protection capabilities when combined with existing software technologies such as volume management and replication.

NetBackup has supported simple disk based data protection using disk Storage Units and VTLs for several releases but with the release of NetBackup 6.5 several new and enhanced disk-based features were introduced, including the following:

- Intelligent disk management capabilities to enable more efficient use of server and storage resources.
- Load balancing across multiple NetBackup Media Servers to improve overall backup performance
- Pre-allocation of disk space to prevent backup failure due to lack of disk space.
- Deduplication features to eliminate redundant data from backup storage.
- Integration with intelligent disk based appliances from leading vendors to allow the user to get optimal value from these solutions.
- Support for 'off host' duplication with VTLs to provide tracking within NetBackup for backup duplication carried out within the VTL.

The value of these new and enhanced capabilities is much more than the sum of their individual contributions. The real benefit is in the way they are combined together to build powerful solutions to meet the challenges of protecting ever increasing amounts of data in IT environments where strained IT budgets, high levels of complexity, and strict compliance regulations are the norm.

There are six different types of disk storage devices supported by NetBackup 6.5; including two disk types that together form the Flexible Disk Option.

Disk types available prior to the 6.5 release:

- **BasicDisk** – simple backup to a directory path on a single disk volume with the option of 'first in first out' staging to tape.
- **Virtual Tape Library (VTL)** – a disk storage device that uses a software layer to emulate a tape library. VTLs are an increasingly popular replacement for physical tape libraries as they offer simplified management and reduced operational costs.

New disk types introduced in the 6.5 release:

- **Flexible Disk Option:**
  - **AdvancedDisk** – multiple disk volumes presented as a single storage pool on a Media Server providing dynamic storage allocation.
  - **SharedDisk** – array based disk volumes presented as a storage pool to multiple Media Servers to provide high performance, load balanced, fault tolerant backup storage.
- **PureDisk Deduplication Option** – de-duplicated disk storage leveraging commodity disk.
- **OpenStorage Option** – API level integration with 3<sup>rd</sup> party intelligent disk storage appliances.

This paper explains each of these disk options and their capabilities as well as the other new features in NetBackup 6.5 that make up the Enterprise Disk Foundation. The paper also provides guidance in choosing the appropriate disk options for your environment.

## 1.1 Glossary

The following descriptive terms are used throughout this document:

**Disk based data protection** – the process by which backups are written to, and stored on, disk based data storage rather than tape media as is common in traditional backup models.

**Storage Unit** – a logical target to which NetBackup writes backup data. Storage Units may map to either disk or tape storage. The precise nature of the mapping depends on the type of Storage Unit.

**Storage Unit Group** – a collection of Storage Units that are addressed as a single entity by backup and duplication jobs. Individual Storage Units within the Storage Unit Group that are used for particular operations depend of the rules associated with the Storage Unit Group.

**Disk pool** – a collection of disk volumes that are presented as a single pool of storage that can be used by one or more Storage Units

**SharedDisk disk pool** – collection of disk volumes assembled in a disk pool that can be accessed by multiple Media Servers. A SharedDisk disk pool usually resides on a single disk array and the disks within the disk pool can be presented to any of the Media Servers sharing it.

**Flexible Disk Option** – comprises the AdvancedDisk and SharedDisk disk types, which make use of directly presented disk storage (e.g. a volume on a NAS device or SAN device available to NetBackup). . When licensing either AdvancedDisk or SharedDisk the same, Flexible Disk Option, license is used.

**Intelligent storage appliance** – a device that provides dedicated, intelligently managed storage that is more than just as simple disk array. In this paper the term 'storage appliance' is applied to intelligent devices that store backup data.

**Deduplication** – the process of eliminating redundant data stored for backup and recovery purposes. Deduplication generally involves breaking down the data into segments, recording the segments for tracking purposes and writing unique segments to storage. Deduplication may take place during the initial write or during post-processing, depending on the device.

**Deduplicating device/appliance** – a storage device in which data is broken down into segments and only one copy of each segment is stored, irrespective of how many files contain that segment. Deduplicating devices offer significant space savings over conventional disk storage but can have slower transfer rates than conventional disk.

**Storage Lifecycle Policy** – a mechanism which controls the creation and retention of multiple copies of the same backup data providing a plan or map of where backup data will be stored and for how long.

**RPO (recovery point objective)** – a service level figure defining the maximum allowable time, prior to a failure, to which an application must be recovered. For example, a recovery point objective of 24 hours means recovered data should no more than 24 hours older than the data at the time of failure. Actual recovery points may differ from established objectives.

**RTO (recovery time objective)** – a service level figure defining the maximum time allowed to recover an application, server or series of application or servers following a failure. Again it should be noted that the actual recovery time may differ from any established objective. Proper planning and testing needs to be carried out in order to assure that the actual recovery time aligns with the desired recovery time objective.

## 2.0 NetBackup Enterprise Disk Foundation

The Enterprise Disk Foundation introduced in NetBackup 6.5 consists of a collection of new disk types and features designed to optimize the way backups are created using these disk types. Figure 1 below shows the four disk storage types that make up the Enterprise Disk Foundation and some of the key features of each of them together with some common features which apply to all disk types. These features are explored on more detail in the subsequent parts of this section.

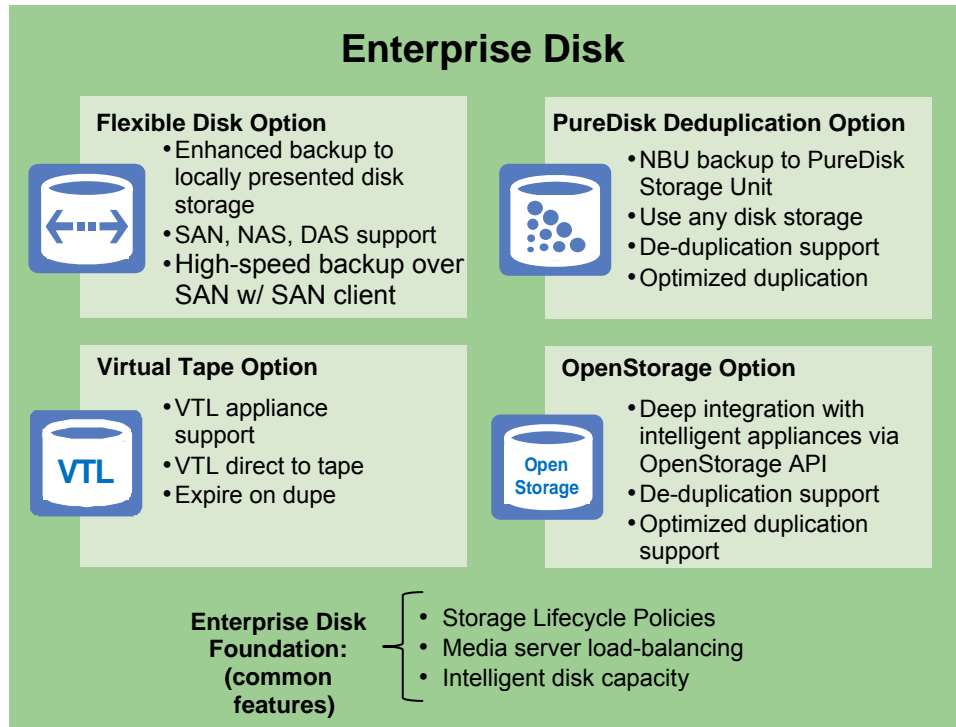


Figure 1 – Overview of the Enterprise Disk Foundation Features

### 2.1 Media Load Server Balancing and Failover

Media Server Load Balancing extends the concept of Storage Unit Groups, originally introduced in NetBackup 5.1, to create a proactive system for routing client backups to the least heavily-loaded Media Server in a Storage Unit Group. Media Server Load Balancing ensures that no single Media Server becomes a bottleneck and avoids the single point of failure of a 'downed' Media Server.

Media Server Load Balancing and Failover increases server utilization and availability by enabling NetBackup Media Servers to be grouped so that backup, restore, and duplication jobs are load-balanced across the group of Media Servers automatically and dynamically. The algorithm for selecting the best Media Server will assess the CPU and memory utilization, and number of jobs currently in progress on each Media Server to select the "best" Media Server for the next NetBackup job. If a Media Server fails for any reason, new jobs will be assigned to other Media Servers in its group until the failed Media Server is repaired so no new backup jobs will fail. Where checkpoint restart is possible, backup jobs running at the time of a Media Server failure are check pointed and will resume automatically on another Media Server within the same group.

The 'grouping' of Media Servers is achieved in one of two ways:

1. Media Servers are defined as associated with a pool of disks (this may be either as 'members' of a disk pool or by means of access credentials for a storage appliance).
2. Media Servers are associated with a Storage Unit Group with the "load balance" radio button checked in the definition configured through the NetBackup Administration GUI.

Media Server Load Balancing looks at all Media Servers in a Storage Unit Group or disk pool and directs the backup to the least heavily loaded server. It uses a load checking algorithm that measures CPU and memory utilization on each Media Server (ranking the server from 0, heavily loaded, to 3, idle) and the number of jobs running on the server.

## 2.2 Intelligent Disk Capacity Management

Intelligent Disk Capacity Management increases backup job success rates by automatically selecting an available disk volume for the backup target based on available capacity, estimated job size and the pre-allocating or "committing" the space on that volume so that other backup jobs do not attempt to use the same space.

NetBackup places a priority on writing the backup image to a single volume if possible; if no volume is large enough to store the entire backup image, multiple volumes can be spanned as necessary.

NetBackup will treat a disk as "full" if the sum of both the committed space and the used space exceeds the high-water mark. If a disk is determined to be full with this method, NetBackup will not use the disk until space is released on it. This feature eliminates the majority of backup failures due to full disk conditions.

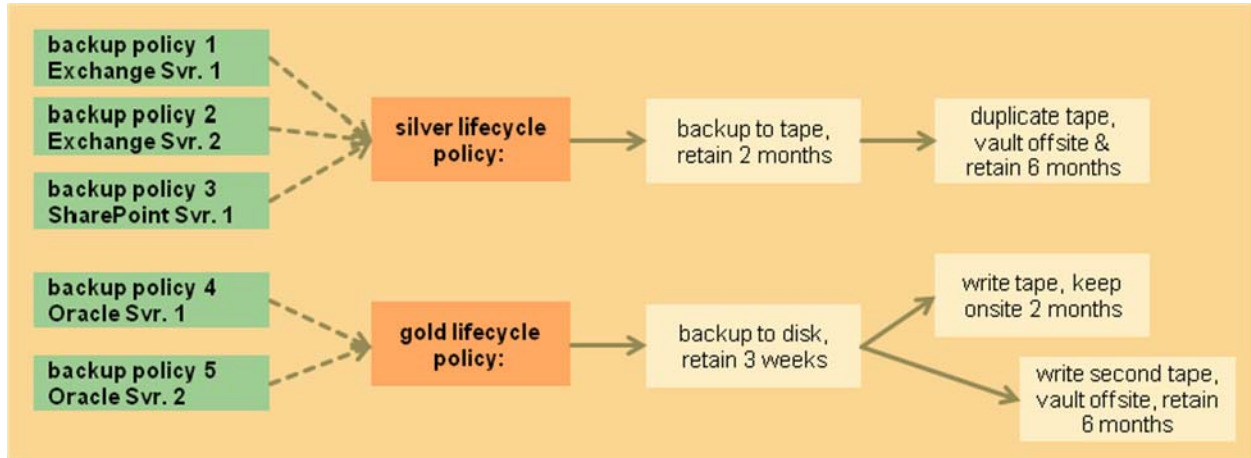
Intelligent Disk Capacity Management only works with the Flexible Disk, OpenStorage, and PureDisk options, which allow the amount of available space to be accurately determined, and cannot be used with BasicDisk Storage Units or with VTLs (which appear to NetBackup as tape).

## 2.3 Storage Lifecycle Policies (Lifecycles)

A Storage Lifecycle Policy allows administrators to automatically specify where backup data will be stored and for how long. The Storage Lifecycle Policy determines the locations where the backup is initially written and the destinations where it is subsequently duplicated to as well as the period of time that each copy of the backup will be retained. It also automates the duplication process making Storage Lifecycle Policies an ideal technology for implementing standardized protection policies.

A Storage Lifecycle Policy primarily consists of two core components: a list of storage destinations where copies of the backup images will be stored, and the retention period for each copy. A Storage Lifecycle Policy can replace both the conventional duplication and staging processes with a single automated processing sequence by introducing a series of storage locations, or destinations, using different types of storage with different retention periods. Storage Lifecycle Policies ensure backup data always exists at the appropriate locations at the appropriate phases of the lifecycle.

A Storage Lifecycle Policy is reusable by many NetBackup Backup Policies. If a storage plan changes (e.g. a new regulation is imposed on your business requiring changes to retention periods or the number of copies created), you simply need to change a small number of Storage Lifecycle Policies and all of the backups will automatically employ the changes.



**Figure 2 - Retentions within Storage Lifecycle Policies**

Figure 2 above shows how two Storage Lifecycle Policies can be applied to several backup policies to ensure that the correct number of copies of each backup is created and retained for the correct period of time.

Storage Lifecycle Policies do not migrate data from one location to another progressively over a period of time. Instead, they ensure that all the image copies are created as soon as possible and apply different retention periods to each copy. Copies held on higher cost, higher performance storage typically have shorter retention periods than copies held on lower cost, slower access storage and space on this premium storage is released sooner. After the original backup completes, the Storage Lifecycle Policy process creates copies of the image, retrying as necessary to ensure that all copies are successfully created.

### 2.3.1 Capacity Managed Retention and Staging

Storage Lifecycle Policies allow the use of a new feature called Capacity Managed Retention. Capacity Managed Retention has been introduced to support smart cleanup of backup images staged to disk pools. With Capacity Managed Retention, the retention period is the time the image should ideally be kept on the storage device. This time period is known as the “try-to-keep time”. NetBackup will keep the image copy until such time as the space is required for other backups. This may be a longer or shorter period than the try-to-keep time depending on the demand for space. Images are expired and removed from disk when the disk high water mark is met with the less critical images being removed first, starting with the oldest ones, which allows the more important images to be retained as long as possible.

Capacity Managed Retention can be applied to all disk based backup copies in a Storage Lifecycle Policy, however the longest retention period used by any copy must be fixed rather than capacity managed. Consequently, if copies are only created on disk, at least one copy must have fixed retention and that retention period must be longer than any of the capacity managed retention periods. Copies of backups written to tape must have fixed retention periods and these must also be longer than any capacity managed retention periods defined for disk based copies. If a storage device cannot be configured in NetBackup to have high and low water marks (e.g. VTLs or other devices that present disk storage as an emulation of tape storage) it cannot be used with capacity-based retention.

### 2.3.2 Data Classifications

Some types of data are more important than others (for example, medical records or financial data are probably more important than the pictures from the office party) and these types of data often have tighter RPO and RTO requirements than less critical data. Data Classifications have been introduced in NetBackup 6.5 as a mechanism for classifying backups by importance. Data Classifications can be applied to both Backup Policies and Storage Lifecycle Policies and, when used in conjunction with capacity managed retention, enable administrators to get higher utilization out of their storage hardware while providing potentially faster restore times on their most important data.



# White Paper: NetBackup™ 6.5: NetBackup Disk Based Data Protection Options

Data Classifications are used by Storage Lifecycle Policies to apply a rank to backup images written to the same capacity managed Storage Units. Backup Images are then retained for different periods of time based on their importance, overriding the traditional ‘first in first out’ model associated with Disk Staging Storage Units. Data Classifications ensure that the more critical data (as identified by administrators) remains on the fast recovery storage hardware in preference to less critical data that may be sharing the same storage.

Data Classifications can also be used in NetBackup 6.5 as a simple way of identifying, grouping and reporting on Backup Policies with a common level of importance. Data Classifications can be used even where Storage Lifecycle policies are not implemented. (Reports based on Data Classifications are available through Veritas Backup Reporter.)

Storage Lifecycle Policies and data Classifications are discussed in more detail in a separate whitepaper titled “Implementing Storage Lifecycle Policies”.

## 2.4 Disk Pools

Disk pools form one of the key concepts underlying the Flexible Disk Option. It significantly changes the way in which available disk space is used. A disk pool groups a set of disks together to form a single block of storage that can be shared among multiple Storage Units and, in some cases, multiple Media Servers.

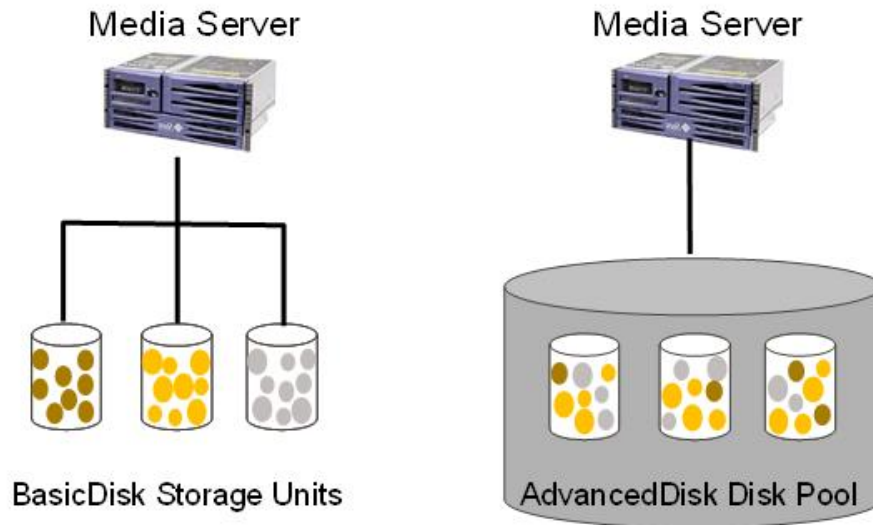
With the Flexible Disk Option disk types of AdvancedDisk and SharedDisk, the disk pool provides a pool of storage for use by the Storage Units, replacing the more conventional one-to-one mapping between disk and Storage Unit provided by the BasicDisk model. In BasicDisk configurations each Storage Unit has access to a single disk volume or part thereof. With the Flexible Disk Option, the Storage Units can access all the disks in a disk pool and the disk used for a particular backup is selected based on the amount of space available. In effect the entire disk pool appears as a single disk to reach Storage Units.

The following table summarizes the advantages of using disk pools over BasicDisk volumes.

Feature	BasicDisk	Disk Pools	Benefit of Disk Pools
Storage sharing	Fixed capacity based on a single disk. Multiple Storage Units can be created on a Single Media Server but each one requires a dedicated disk.	Multiple Storage Units on the same Media Server can share a disk pool. Space is allocated dynamically as required.	Automatic space allocation reduces the risk of backup failure.
Intelligent capacity management	No checks on available space before backup starts	“Will Fit” selection selects the Storage Unit based on ‘available’ space.	Pre-selection of storage ensures efficient backup operation.
Simplified growth	No ability to increase the size of individual Storage Units when disks fill up	Disk pool capacity is increased by adding additional disk volumes into the pool.	Minimal reconfiguration required as environment grows. Improved utilization of disk.
Media Server load balancing	Does not support load balancing in Storage Unit Groups	Supports load balancing – directing the backup to the least busy Media Server	Eliminates Media Server bottlenecks to increase the performance and success of backup jobs.
Storage Lifecycle Policies	Staging achieved using individual Storage Units with staging schedules	Data can be classified and managed differently throughout its life based on the importance of the data. After it has been duplicated, data can be selectively expired based on the relative importance to the business.	‘Important’ backups are available for rapid restore for longer periods, thereby improving RTOs.

**Table 1 - Disk Pools vs. BasicDisk Storage**

Figure 3 below shows the difference between a Media Server with access to BasicDisk Storage Units and a Media Server with access to an AdvancedDisk disk pool.



**Figure 3 - BasicDisk vs. AdvancedDisk**

In the BasicDisk configuration the three disks act as separated Storage Units and the backups are written to specific disks. Backup jobs do not use all three disks unless they are configured to form a Storage Unit Group and there is no guarantee that the disk space is used efficiently among the three independent disks.

AdvancedDisk utilizes space more efficiently than BasicDisk. In the AdvancedDisk configuration the three disks form a single pool and backups may be written to any of the disks. When a backup runs, Intelligent Disk Capacity Management will automatically estimate the size of the backup and select the disk with the greatest amount of free space and will reserve space to match the estimated size of the backup image so that other backup jobs do not over commit beyond the disk's available space. Note how the different colored backup images reside on different disks in the BasicDisk configuration but are distributed across all of the disks in the AdvancedDisk disk pool.

Storage Units that use AdvancedDisk disk pools on different Media Servers can be configured to form 'load balanced' Storage Unit Groups. When a backup is run using the Storage Unit Group, the Storage Unit on the least busy Media Server is automatically selected and a suitable disk within the AdvancedDisk disk pool on that Media Server is used for the backup.

An AdvancedDisk disk pool can be formed from any disks that can be seen by, and are mounted to, the Media Server. However it is recommended that all the disks in a particular disk pool should have common characteristics to ensure consistent behavior when different disk are selected for use.

In a SharedDisk configuration the disk pool is created using a disk array which can be shared between multiple Media Servers over a SAN. This allows the individual disks within the disk pool to be presented to different Media Servers at different times. As in the case of AdvancedDisk, the disk selection process is based on the 'will fit' model with space being reserved while a backup is in progress to prevent conflicts but with SharedDisk the disks can be dismounted from one Media Server and then mounted against another Media Server.

SharedDisk disk pools load balance at the Storage Unit level rather than the Media Server. Multiple Media Servers have access to the same Storage Unit. When a backup is run to a particular Storage Unit, the least busy Media Server is automatically selected and a suitable disk is mounted on that Media Server for the backup to write the backup image. Figure 4 shows four Media Servers sharing a common SharedDisk disk pool.

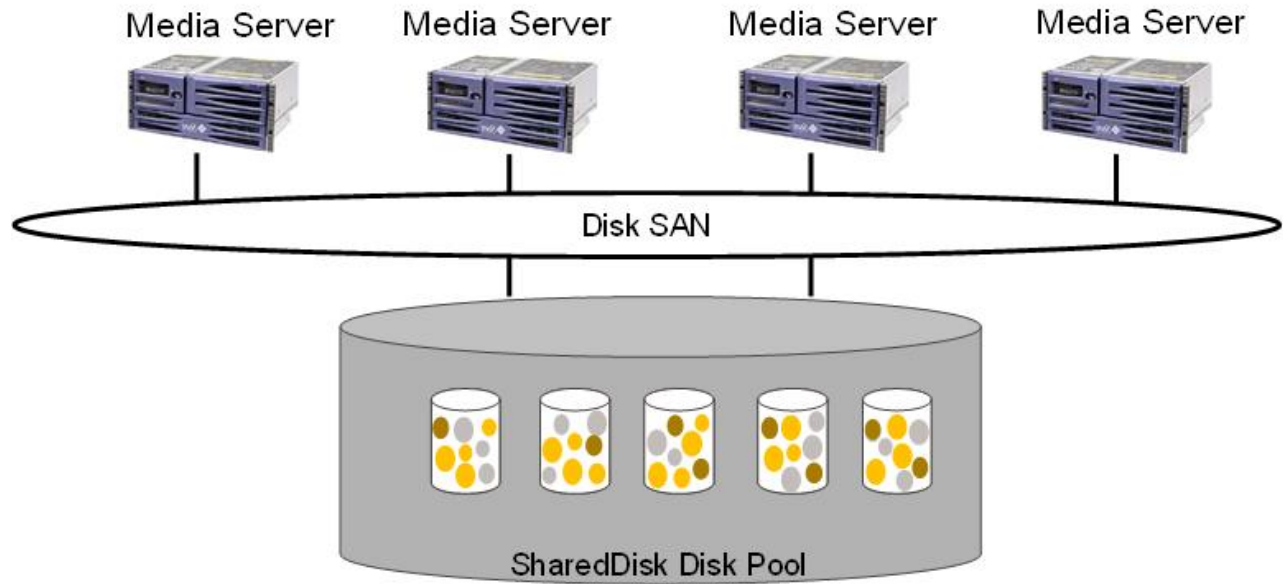


Figure 4 - SharedDisk Disk Pool

## 2.5 Intelligent Disk Storage Devices

An intelligent disk storage device combines disk storage with software and is presented to NetBackup Media Servers as a device rather than a simple collection of disks. Prior to the release of NetBackup 6.0 the only intelligent disk storage devices that could be used with NetBackup were VTLs. NetBackup 6.0 introduced support for the NetApp Nearstore device which provides space efficient deduplicating disk storage based on the NetApp WAFL file system. NetBackup 6.5 extends this capability through integration with PureDisk and the OpenStorage API which allows the use of a wide range of OpenStorage compliant devices. OpenStorage compliant devices generally provide deduplicating capabilities and other features such as device level replication and energy efficient storage. These 'second generation' intelligent disk devices can offer significant advantages over VTLs because they present the disk storage as disk rather than emulating tape storage and can thus make full use of the Enterprise disk Foundation features. Table 2 below summarizes the advantages of these intelligent disk devices over VTLs.

Feature	VTL	Intelligent Disk Device	Benefit
Parallel processing	Only one read or write operation allowed at a time per virtual tape significantly extending the time required to effectively protect data.	Concurrent read or write operations can occur per volume potentially completing all jobs in a much shorter time.	Shorter backup windows.
Resource sharing	Only one Media Server can read or write to a virtual tape at any given time	Multiple Media Servers can read or write to the same volume concurrently.	More efficient resource sharing between Media Servers.
Granularity – duplication	The virtual tape is smallest unit of capacity. Images written to a virtual tape cannot be copied until all read or write operations to the virtual tape are complete.	The backup image is the unit of capacity. Images on a disk volume can be copied while other images on the same volume are being read or written.	More efficient duplication cycles.

Feature	VTL	Intelligent Disk Device	Benefit
Granularity – expiration	The virtual tape is smallest unit of capacity. All images on the virtual tape must expire before any space is released.	The backup image is unit of capacity. Free space is available as soon as a backup image is expired.	More efficient utilization of storage.
Duplication	NetBackup is unaware of any replication that takes place at the VTL level.	Optimized duplication ensures NetBackup is aware of replication activity.	Copies created by replication can be used for restore by NetBackup.
Capacity Management – selection	Virtual tapes must be filled in sequence.	Volumes can be selected based on available space.	Intelligent capacity management for better storage utilization.
Capacity Management – staging	Virtual tapes cannot be used for disk staging as NetBackup does not know how full each virtual tape is.	Volume high and low watermarks allow staging operations.	Capacity managed expiration.

**Table 2 - VTL vs. Intelligent Disk Storage**

## 3.0 NetBackup Disk Types

This section examines each of the six disk types identified in the introduction, explaining the key features of each type and their relative strengths and weaknesses.

### 3.1 BasicDisk

#### 3.1.1 What is BasicDisk?

BasicDisk is any simple disk volume that is made available to NetBackup as a target. It could be a mounted volume from a NAS share, direct attached storage within a NetBackup Media Server (e.g. the C:\ drive) or a SAN attached volume.

BasicDisk Storage Units can be configured as 'staging' areas from which backups are duplicated to alternative storage but this process is limited to a 'first in first out' behavior requiring individual Storage Units to be configured for backups of differing importance.

#### 3.1.2 What's good about BasicDisk?

- **Low capital cost** – BasicDisk is included in the standard NetBackup server license and standard infrastructure license and can also be used with low-end disk so it has the lowest purchase costs of any disk type. However the absence of intelligent automated management features means that the total cost of ownership of BasicDisk may be higher than some of the other disk options discussed in this paper.
- **Wide hardware support** – BasicDisk can be used with any disk storage that can be presented to a Media Server with a file system that can be written to, irrespective of the connection method or disk characteristics.
- **Alternate Media Server restore and content replication** – BasicDisk volumes can be manually presented to alternate Media Servers if restores are required when the host Media Server is down. The contents of a BasicDisk volume can also be easily replicated to a DR site and presented to a Media Server there for disaster recovery purposes.

#### 3.1.3 What are the downsides of BasicDisk?

- **Increased management overhead:**
  - BasicDisk does not support Enterprise Disk Foundation features such as Storage Lifecycle Policies, Intelligent Disk Capacity Management and Media Server Load Balancing which provide automated resource management
  - Unlike other disk types described in this section, BasicDisk Storage Units are confined to the individual disk volumes. BasicDisk storage is not presented or managed as pools of disk volume where extra capacity can be added by simply adding additional volumes into the pool. Instead, creating extra capacity requires configuring additional BasicDisk Storage Units and changing backup policy settings or Storage Unit Group policies.
- **Staging limitations** – while BasicDisk storage can be used as a 'staging' area prior to writing to tape, BasicDisk Storage Units operate on a 'first in first out' model and do not support the more advanced staging through lifecycles.

#### 3.1.4 When should I consider using BasicDisk?

BasicDisk can be used for simple disk staging or smaller disk backup environments but may not be appropriate for wide spread use in larger environments where performance, resource utilization, automation, and high availability are important. BasicDisk is especially well suited to support NetBackup catalog backups due to the simplicity of the disk storage presentation to NetBackup, which simply uses a server name and a path.

NetBackup's EMM database can be easily rebuilt (or remapped to see the replicated volume instead) in order to restore the catalog.

The lack of support for Enterprise Disk Foundation features limits the capabilities of BasicDisk. BasicDisk should not be used in any configuration where load-balancing across servers and disk volumes is required or where advanced staging capabilities such as Capacity Managed Retention are needed.

### 3.2 AdvancedDisk

#### 3.2.1 What is AdvancedDisk?

AdvancedDisk increases the utilization and performance of disk by combining disk resources together as a single pool rather than treating them as smaller, separate islands of disk. AdvancedDisk also extends the capability of BasicDisk through the Enterprise Disk Foundation features which increase Media Server utilization and availability while improving backup success rates and performance. Instead of creating Storage Units from single disk volumes, AdvancedDisk allows you create Storage Units containing multiple disk volumes defined within a disk pool.

#### 3.2.2 What are the benefits of AdvancedDisk?

- **Enterprise Disk Foundation feature support** – AdvancedDisk can make use of features such as intelligent capacity management, Storage Lifecycle Policies and Media Server Load Balancing
- **Wide hardware support** – like BasicDisk, AdvancedDisk can support most types of disk storage presented to a Media Server.
- **Easy upgrade path from BasicDisk** – AdvancedDisk uses the same hardware as BasicDisk and BasicDisk Storage Units can easily be upgraded to AdvancedDisk Storage Units provided the Storage Unit already exclusively owns the entire file system.
- **Reduced management overhead** – automated infrastructure reduces the level of administrator attention required by distributing job load and intelligently managing disk capacity to ensure that jobs do not fail because of out of space conditions until all the available disks are full.
- **Dynamic storage allocation** – because the Storage Units access a pool of disks, more capacity on other disks is automatically 'added' to a Storage Unit when a disk fills up. The disk pool model also makes it simple to increase overall storage capacity as additional disks can be added to the disk pool without needing to reconfigure the Storage Units which use it.

#### 3.2.3 What are the downsides of AdvancedDisk?

- **Must use local server for restore and duplication** – AdvancedDisk storage cannot easily be presented to an alternative Media Server to allow restore and duplication operations. This can prevent restore operations in the event that the Media Server which wrote the backup is down.
- **No replication capability** – AdvancedDisk storage cannot easily be replicated to a DR site and mounted on a DR server because it uses a disk pool instead of individual disks and mount points.

#### 3.2.4 When should I consider using AdvancedDisk?

AdvancedDisk is well suited for use where fast and flexible backup to disk is required, especially for short term staging prior to writing data off to other storage media. AdvancedDisk can work with "whatever disk you have laying around" and assemble it into a highly utilized disk target it. This makes it an attractive option for customers that may have spare disk storage which is no longer required for primary storage and can be re-purposed as backup storage. The load balancing and capacity management features of Enterprise disk Foundation mean that AdvancedDisk offers clear advantages over BasicDisk in terms of more efficient disk and server utilization.

## 3.3 SharedDisk

### 3.3.1 What is SharedDisk?

SharedDisk allows LUNs in a disk array to be shared between multiple Media Servers in the same way that tape drives can be shared with the Shared Storage Option. As in AdvancedDisk, the SharedDisk volumes can be defined in a disk pool, however in SharedDisk storage the disks are in a storage array and the disk pool is presented over a SAN connection to multiple Media Servers.

Disks in a SharedDisk disk pool are not presented to multiple Media Servers simultaneously; instead they are dynamically assigned to individual Media Servers when required. When a disk from the pool is assigned to a Media Server, the file system is mounted, and the backup image is then read or written. Once the backup or restore operation has completed the disk is dismounted and returned to the pool. Any disk within the disk pool may be presented to any Media Server at any time. Since a disk can be mounted on different Media Servers, the Media Servers participating in the pool must be the same platform type running the same operating system.

### 3.3.2 What's good about SharedDisk?

- **Superior scalability** – offers ability to scale in two dimensions by adding storage (additional LUNs) or processing power (additional Media Servers).
- **Enterprise Disk Foundation feature support** – SharedDisk can make use of features such as intelligent capacity management, Storage Lifecycle Policies and Media Server Load Balancing
- **Improved capacity utilization** – as the disks in the disk pool can be used by multiple Storage Units and presented to multiple Media Servers the available disk space can be more efficiently used. Large backups going to one Media Server/Storage Unit combination can share a disk with smaller backups going to another Media Server/Storage Unit combination.
- **High performance** – as the data is written over a direct I/O path to disk rather than requiring additional LAN hops to NAS appliances or SAN hops through storage appliances to the actual disk.
- **Media Server level failure tolerance** – as SharedDisk disks can be presented to multiple Media Servers, backups and restores can automatically be re-directed to a different Media Server if the required Media Server is unavailable. This feature is particularly useful from a restore perspective because, unlike BasicDisk and AdvancedDisk, no storage reconfiguration is required in the event of a Media Server failure.

### 3.3.3 What's the downside of SharedDisk?

- **Single server type** – all servers sharing a disk pool must be the same platform type running the same operating system
- **Limited hardware support** – not all disk arrays can support SharedDisk as array compatible code is required on the Media Server. Disk arrays must be qualified for use with SharedDisk.
- **Limited platform type/operating system support** – not all operating systems can support the SharedDisk feature. At the time of writing it is only available for Windows 2003 and Solaris Sparc Media Servers.

### 3.3.4 When should I consider using SharedDisk?

SharedDisk configurations offer distinct advantages in terms of scalability and resilience over other types of disk storage in environments where a single type of Media Server is used. The high performance nature of array based disk storage combined with the ability of the SharedDisk environment to adapt to server availability makes SharedDisk the ideal choice for the short term storage of applications with very high RPO and RTO requirements.

## 3.4 Virtual Tape Library (VTL)

### 3.4.1 What is a VTL?

VTLs use disk arrays to emulate the behavior of tape drives and tapes in tape libraries using 'virtual tapes'. Their main advantages are that they are more reliable, require less user intervention and occupy less space than their physical tape counterparts. One reason they are more reliable is that virtual tape devices have no minimum streaming speed so they do not suffer damage as a result of slow write rates in the way that physical tape devices do. VTLs eliminate the management problems surrounding physical tape handling, thereby improving the reliability of backups.

VTLs are easy to operate as they behave like physical tape libraries assigning and deassigning space in blocks equivalent to the capacity of a virtual tapes. Virtual tapes appear to operations staff exactly the same as the tapes they are familiar with. Many VTLs also offer capabilities to replicate data or write it to physical tape but as this usually happens outside of NetBackup control, the resulting copy is not known to, and often not directly accessible by, NetBackup.

### 3.4.2 What's good about a VTL?

- **Ease of integration** – as VTLs appear exactly like tape libraries in NetBackup there are no requirements to change operational practices.
- **Platform independent sharing** – VTLs are generally platform independent and can be connected to multiple Media Servers of different types.
- **High performance with parallel write** – most VTLs are optimized for writing, allow large numbers of dedicated 'virtual tape drives' to be configured to allow fast writing to multiple virtual tapes in parallel. In practice, restrictions on the read rate of the source data means that these high transfer rates cannot be achieved.
- **Intelligent storage options** – many VTLs offer additional capabilities such as deduplication, write through to physical tape and replication between VTLs.
- **Limited Enterprise Disk Foundation feature support** – VTLs can make use of Storage Lifecycle Policies and Media Server Load Balancing but do not support any form of intelligent capacity management.

### 3.4.3 What are the downsides of a VTL?

- **Inefficient use of storage space** – virtual tapes suffer from the same limitation as physical tapes, until the backups on a virtual tape have expired none of the space is released. Even if a virtual tape has potential space available (the images have expired), the space cannot be re-used due to the constraints imposed by sequential tape access, all the images must expire before the virtual tape can be 'scratched' and re-used.
- **No intelligent capacity management** – as VTLs emulate tapes they cannot take advantage of Enterprise Disk Foundation's ability to direct backups to the disk with the most available space. VTLs follow the same rules as physical tapes and backups are appended to partially full 'virtual tapes' without checking how much free space remains on at tape.
- **Restricted access to tapes and resident images** – pre-written images are inaccessible while tape is in use, blocking access for restores and delaying duplication until backups complete. Parallel access is not possible – only one image can be accessed at a time, consequently holding all other images hostage.
- **Content replication is not always tracked in NetBackup** – if virtual tape is replicated to a physical tape or another VTL using the VTLs internal functions, the replicated copy is not known to NetBackup and can only be use if the original copy is deleted or masked so that it is not visible. NetBackup also has no ability to limit which images are replicated should you wish to conserve space and follow SLAs where only a subset of the most important images need to be copied. Some VTLs make use of embedded NetBackup Media Servers or the NetBackup 'VTL-direct-to-tape' feature to avoid this problem by duplicating to virtual



or physical tapes with different media identifiers. In these cases NetBackup is aware that the duplication has taken place even though the operation did not involve a Media Server directly.

### 3.4.4 When should I consider using a VTL?

VTLs represent an easy option when moving from a tape based solution to a disk based solution as they behave in the same way as tapes. VTLs make a good choice for improving backup performance and reliability when there is a limited ability to architect a true disk backup solution. Customer's looking to maximize performance and efficiency that have the flexibility to modify their architecture may consider other disk options.

## 3.5 OpenStorage

### 3.5.1 What is OpenStorage?

The OpenStorage initiative is an exciting disk-based innovation which allows NetBackup to take advantage of the new technological advances of intelligent storage devices, including: storage reduction, backup image duplication, synthetic backups, replication and energy efficiency. The tight integration with third party intelligent storage devices offers NetBackup users improved management of their backup images and the added functionality of third party intelligent storage devices without the limitations of tape emulation.

Symantec has developed the OpenStorage API that lets intelligent disk storage appliance vendors write plug-ins to their appliance for NetBackup Media Servers. This provides NetBackup with visibility into the properties and capabilities of the appliance, and control of the backup images stored in the appliance. The appliance storage can then be treated by NetBackup as disk devices rather than tape devices, as in the case of VTLs. Through the OpenStorage API, NetBackup controls when backup images are created, duplicated, and deleted. The OpenStorage appliances control how the images are stored in and copied between appliances. In this manner, the OpenStorage vendors add unique business value to the overall solution via specialized technological innovations such as backup image deduplication, WAN-optimized backup image replication for disaster recovery purposes, power management, and a wealth of other innovations.

With OpenStorage customers can utilize a variety of qualified, best-of breed, disk-based solutions sold by a number of appliance vendors. As with VTLs, customers benefit by having an easy-to-deploy, purpose-built, disk-based protection solution. Unlike virtual tape solutions, OpenStorage intelligent disk appliances present disk storage to NetBackup rather than presenting disk as tape. The result is that a number of limitations inherent to tape are avoided.

Integration with NetApp NearStore has been available since the release of NetBackup 6.0 and is based on an earlier implementation of the OpenStorage API than the one implemented in NetBackup 6.5. The NetApp devices that integrate with the newer OpenStorage API, taking advantage of Enterprise Disk Foundation features such as Media Server Load Balancing that are not available in the existing NearStore integration, are expected to be available in 2008. See the Symantec Support website for an up-to-date list of supported OpenStorage vendors on the NetBackup Hardware Compatibility List.

### 3.5.2 What's good about OpenStorage?

- **Enterprise Disk Foundation feature support** – OpenStorage appliances can make use of features such as intelligent capacity management, Storage Lifecycle Policies and Media Server Load Balancing
- **Leverages the capabilities of the OpenStorage appliance** – the capabilities vary between different appliances.
  - **Reduced storage requirements** – appliances that support deduplication require significantly less back end storage than the conventional disk types supported by BasicDisk and the Flexible Disk Option.
  - **Deduplicated replication for improved disaster recovery** – NetBackup can take advantage of the replication capabilities of the OpenStorage appliances to “off-host” duplicate NetBackup images to secondary appliances. Space optimized replication capabilities allow ‘off-site over the wire’ capability.

The cost of bandwidth, tape shipping, and storage associated with vaulting data to another data center is drastically reduced.

- **Energy/power savings** – when used with appliances that support energy saving MAID (massive array of idle disk) technologies.
- **Platform independent storage sharing** – the OpenStorage appliance can be presented to multiple Media Servers running on different operating systems.
- **Single point of control** – the OpenStorage appliances are controlled through NetBackup. The migration or copying (the storage lifecycle) of the data within the OpenStorage appliance is controlled through NetBackup.

### 3.5.3 What are the downsides of OpenStorage?

- **Deduplicating appliances challenge management predictability:**
  - **Unpredictable storage demands** – space saving within the appliance is only as good as the vendor's deduplicating algorithm and some may not be as efficient as others.
  - **Unpredictable space reclamation** – when backup images are removed from conventional disk storage devices the amount of space freed up is directly proportional to the size of the backup image. For deduplicating devices there is no proportional relationship between the backup image and space occupied. This problem is compounded by some deduplicating devices that do not reclaim individual unique blocks as they are released. For this reason deduplicating devices are not suitable for use as 'staging' areas and may 'under report' available space to NetBackup's Intelligent Disk Capacity Management system.
- **Unsuitable for staging use** – the relatively low read speed of deduplicating appliances (a side effect of 'reconstituting' the backup from the deduplication segments) means that they may not be suitable for 'staging' areas or as source copies for duplication to tape or high speed disk storage.

### 3.5.4 When should I consider using OpenStorage?

OpenStorage appliances offer various benefits over conventional disk devices including deduplication, optimized duplication and reduced power consumption. They also offer many benefits over VTLs in terms of space utilization and operational efficiency. The decision on when to use an OpenStorage appliance, and which OpenStorage appliance to use, will depend on the user requirements.

## 3.6 PureDisk Deduplication Option

### 3.6.1 What is PureDisk Deduplication Option?

The PureDisk Deduplication Option (PDDO) allows a Media Server to store ordinary NetBackup data in a PureDisk Storage Pool leveraging PureDisk global deduplication technology. As with OpenStorage, NetBackup maintains visibility and control of the replication and expiration process and thus the entire lifecycle of the backup image.

NetBackup clients can transmit backup data over the LAN or SAN to a collection of NetBackup Media Servers. These Media Servers then deduplicate the backup stream before storing it to the PureDisk 6.5 storage server. Existing or new PureDisk clients can also backup to the PureDisk Storage Pool.

### 3.6.2 What's good about PureDisk Deduplication Option?

- **Global deduplicated storage integrated directly into NetBackup** – With deduplication factors typically between 10 and 50, PDDO provides significant savings in storage costs. Its integration directly into NetBackup brings ease of management and a very low cost of ownership.
- **PureDisk capabilities for all NetBackup clients** – Any existing NetBackup client or NetBackup agent can store its data on the PureDisk storage server through the NetBackup Media Servers.

- **Enterprise Disk Foundation feature support** – PureDisk can make use of Storage Lifecycle Policies and Media Server Load Balancing. Intelligent capacity management does not apply to PureDisk appliances as the entire storage pool appears as a single volume to NetBackup.
- **Encryption built in** – PDDO can automatically encrypt all data stored on in the PureDisk Storage Pool.
- **Central point of control** – The client backup data stored in the PureDisk storage server is fully managed by NetBackup.
- **Deduplicated replication** – NetBackup can take advantage of the WAN-optimized replication capabilities of PureDisk Storage Pools to “off-host” duplicate NetBackup images to secondary storage pools. Duplicated images can be stored with different retention levels, treated independently from the original copy, and be verified and cataloged by NetBackup.
- **Scalable performance and capacity** – A PureDisk Storage Pool is a modular system that can be scaled in performance and capacity by adding ‘content router’ nodes. Global deduplication happens across all content router nodes in the storage pool, maintaining deduplication efficiency even when storage nodes are added to the system. Data deduplication is performed on the NBU Media Servers. Multiple streams can be deduplicated in parallel on the Media Server and multiple Media Servers can deduplicate data in parallel as global deduplication optimizes the data across the whole dataset.
- **Commodity disk** – PureDisk works with a broad list of disk systems (DAS, SAN, NAS, iSCSI) which makes it easily deployable in an existing storage environment.

### 3.6.3 What are the downsides of PureDisk Deduplication Option?

- **May require additional CPU resources on the NetBackup Media Server** – In order to stream only unique data to the PureDisk system, the Media Server breaks the image into segments and checks to see if segments are currently stored within the storage pool. This means the Media Server requires more CPU power than a similar server backing up to conventional disk. To reduce the CPU impact of deduplication, backup streams can be spread over multiple Media Servers while still maintaining deduplication optimization.
- **Lower throughput than other disk types** – as deduplication takes place ‘in line’ at the time of the backup, the write speed to PDDO is often less than BasicDisk, AdvancedDisk, SharedDisk and other non-deduplication disk types. PDDO can be used in-line for LAN speed backups up to 100MB/s. For SAN rate backups over 100MB/s it is recommended to use a Storage Lifecycle Policy to stage to a non-deduplicated disk device and then duplicate the backup post-process to PDDO for long term retention.
- **Unsuitable for staging use** – the relatively low read speed of PDDO Storage Units (a side effect of ‘reconstituting’ the backup from the deduplication segments) means that they may not be suitable for ‘staging’ areas or as source copies for duplication to tape or high speed disk storage.

### 3.6.4 When should I consider using PureDisk Deduplication Option?

PureDisk Deduplication Option is ideally suited for longer term retention of data on disk because it offers dramatic space savings. Its built in replication also makes it ideal for customers who may look to replace tape warehousing with electronic vaulting to their own data centers or third party service providers. PureDisk is a software solution that allows using your preferred storage hardware. The PureDisk architecture scales in performance and capacity by adding Media Servers, deduplication plug-ins and content router nodes. Customers who are considering VTLs or hardware based deduplication devices may want to consider PureDisk Deduplication Option. PureDisk provides excellent space reduction and optimized duplication through its deduplication and replication capabilities but is generally slower for both backup and restore than conventional disk storage. PDDO can be used in-line for LAN speed backups but will generally slow down SAN backup and restore rates compared to conventional disk backups. In SAN backup environments, PureDisk storage pools should be used in conjunction with regular AdvancedDisk or SharedDisk ‘staging’ devices to provide long term duplicated storage.

## 4.0 Implementing a Disk Based Solution

The previous sections have shown there are now several different disk types available, each of which has its own unique set of capabilities – so which disk type is right for you? In many cases the ‘right’ choice of disk based data protection is a mixture of different disk technologies; AdvancedDisk and SharedDisk for load balancing and performance, OpenStorage and PDDO for space and power efficient long term storage. This section compares the different disk types and how they can be used to best advantage and presents some examples of where combinations of different disk types are used to meet different requirements.

### 4.1 Choosing the Right Disk Type for the Job.

Choosing the right disk type to meet a specific requirement is crucial. There are three factors that will influence the decision process:

- Capability – is the type of disk suitable for the intended purpose?
- Cost of ownership – does this choice represent the best value for money?
- Compatibility – is this choice compatible with my existing infrastructure?

#### 4.1.1 Capability

As we have seen, each disk type has its strong and weak points, which means that some disk types are more suitable than others for particular tasks.

Table 3 summarizes the capabilities of the five main disk types (AdvancedDisk and SharedDisk are treated as a single disk type here titled Flexible Disk) and illustrates their relative strengths in five key areas.

	VTL	OpenStorage	Flexible Disk	PureDisk	BasicDisk
Backup Performance <sup>1</sup>	●●●●○	●●●○○	●●●●○	●●●○○	●●●●○
Recovery Performance <sup>1</sup>	●●●●○	●●●○○	●●●●○	●●○○○	●●●●○
Lifecycle management capabilities <sup>2</sup>	●●●○○	●●●●○	●●●○○	●●●●○	○○○○○
Storage Efficiency <sup>3</sup>	○○○○○	●●●●○	●○○○○	●●●●●	○○○○○
Deduplication capabilities <sup>4</sup>	●○○○○	●●●●●	○○○○○	●●●●●	○○○○○

**Table 3 - Summary of Capabilities**

Note: Both VTL and OpenStorage ratings will be highly dependent on the hardware selection.

1. Backup Performance and Recovery Performance will typically be impacted by deduplication or compression being utilized. If you are using a storage appliance check with your appliance provider for typical performance impact.

## White Paper: NetBackup™ 6.5: NetBackup Disk Based Data Protection Options

2. Some VTLs have “direct to tape” functionality for use with NetBackup which may increase their performance.
3. VTLs must mimic tape cartridges which limits recovering storage space.
4. Deduplication capabilities vary widely with appliance vendors. Check with your appliance provider for typical deduplication capabilities. Flexible disk does not include deduplication and PureDisk does.

The choice of disk storage depends largely on two factors, whether the requirement is for short term or long term storage and whether the solution will use commodity disk or a storage appliance of some type.

Shorter term storage is generally used to provide rapid recovery to the last good backup point in order to meet recovery SLAs. For this reason short term storage is chosen based on its speed and is often used to provide a ‘staging’ capability for data that is then copied to longer term storage. For customers desiring an appliance form factor either VTL or OpenStorage devices can handle very fast movement of data. While OpenStorage provides the same benefits and more compared to VTL, some customers may have existing hardware that does not support the OpenStorage technology making VTL a potential choice. Customers who have the skills to manage commodity disk pools may opt for Flexible Disk to maximize savings and performance.

Longer term storage is generally used to hold data that must be kept in accordance with compliance SLAs. Long term storage is chosen for its capacity and will likely use deduplication to make it economically feasible to hold large amounts of data on disk. For customers who prefer an appliance approach, OpenStorage is preferred because it provides much more robust data management capabilities and potentially enables better deduplication rates both of which increase functionality while decreasing costs. PureDisk brings the same deep integration with NetBackup for robust management but adds significant deduplication and replication capabilities in the software for use with any commodity disk available.

Table 4 shows how these two selection criteria can be applied to determine the most appropriate disk type to be used in each case.

	Short Term (1-14 days)	Longer Term (14-180 days)
Appliance	VTL or OpenStorage Device	Deduplicating VTL or OpenStorage Device
Commodity Disk	Flexible Disk or BasicDisk	PureDisk

**Table 4 - Disk type selection criteria**

This table is intended to be a guide only. You may have special considerations for your environment which would lead you to different conclusions. If you need assistance planning your data protection environment please contact your Symantec account representative for assistance.

### 4.1.2 Cost of ownership

The total cost of ownership of different disk types can be broken down into two components, the capital cost of the initial hardware and the on-going operational cost of running the backup environment.

The capital cost of appliance based solutions is generally higher than that of commodity disk based solutions as additional hardware and software components are required. However it is important when considering deduplicating devices such as PureDisk and OpenStorage appliances to remember that the deduplicating effect means that ‘cost per gigabyte’ of storage may be significantly less than it appears as the actual space used on the device is less than on an equivalent conventional disk or VTL.

Operational costs of deduplicating solutions are also lower as the reduced storage capacity translates into reduced power and cooling requirements. The ‘self managing’ nature of the appliances mean they tend to be more reliable and require less operator intervention than commodity disk solutions and, by the same token, the Enterprise Disk Foundation features mean that Flexible Disk Option implementations also have a lower operating cost than BasicDisk implementations. Thus while a BasicDisk solution is likely to have one of the lowest capital

costs per gigabyte it also likely to have the highest operational cost as there is no automatic management capability.

### 4.1.3 Compatibility

It is important to remember the following points when considering the choice of disk type:

1. Not all disk types are supported with all Media Server platforms and versions of NetBackup 6.5.x.
2. OpenStorage devices and PureDisk are delivered outside of NetBackup's packages and on different release schedules. Integration timescales with NetBackup are dependent on multiple development groups, including third-party vendors.
3. SharedDisk is dependent on disk array support and, although many leading disk arrays are supported, there are some array types that cannot be used.

Reference should be made to the latest NetBackup compatibility lists to ensure that the proposed solution is supported with the existing or proposed infrastructure.

## 4.2 Disk Based Solution Examples

The following examples show how a combination of disk types can be used to meet the short term and long term backup requirements.

### 4.2.1 Balanced backup for optimal throughput

The use of AdvancedDisk and SharedDisk Storage Units with Storage Lifecycle Policies provides an efficient mechanism for making initial backups that can subsequently be duplicated to tape or deduplicating storage. Figure 5 shows clients backing up to a number of Media Servers which are operating in a load balanced configuration with a SharedDisk Storage Unit and a SAN attached OpenStorage appliance. The Storage Lifecycle Policy ensures that the backups are written to the SharedDisk storage and then duplicated to the OpenStorage appliance for long term storage.

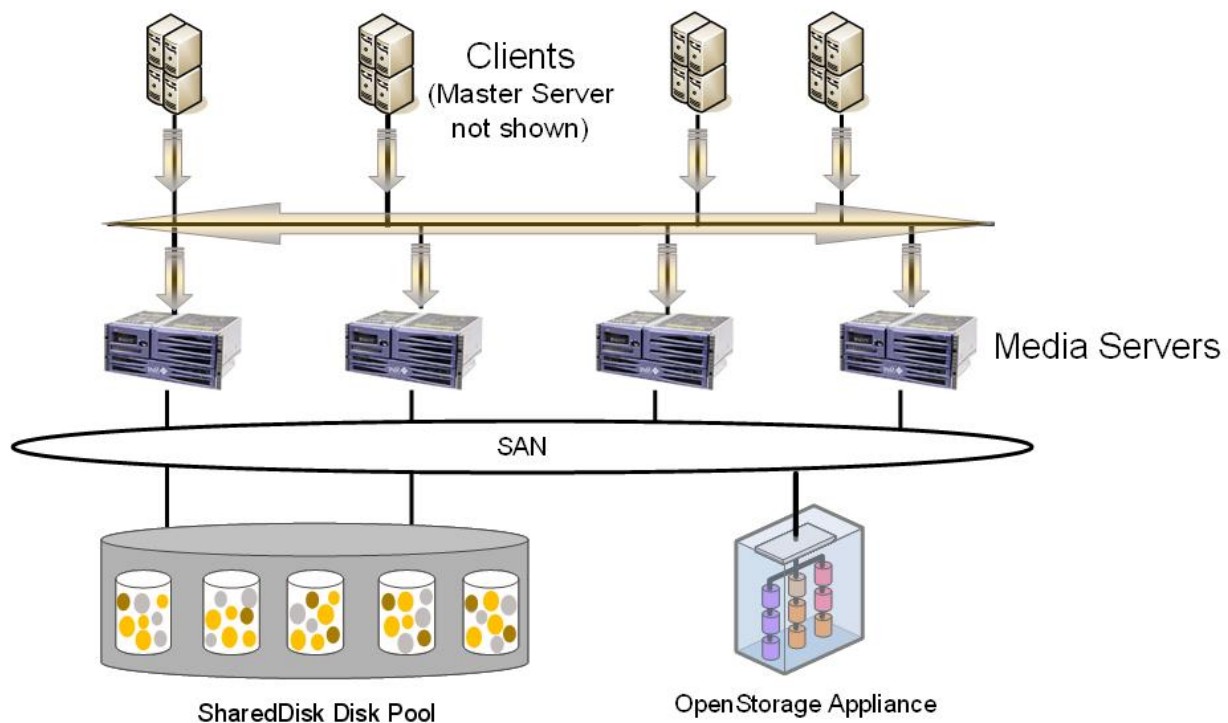


Figure 5 - Load balancing with SharedDisk and OpenStorage

The down side of this configuration is that, because SharedDisk is used, it requires all the Media Servers to be the same platform. Figure 6 below shows a similar configuration using AdvancedDisk storage with a LAN attached OpenStorage appliance. In this case Storage Unit Groups are used to achieve the initial load balancing across the Media Servers and Storage Lifecycle Policies control the duplication of backups to the OpenStorage appliance.

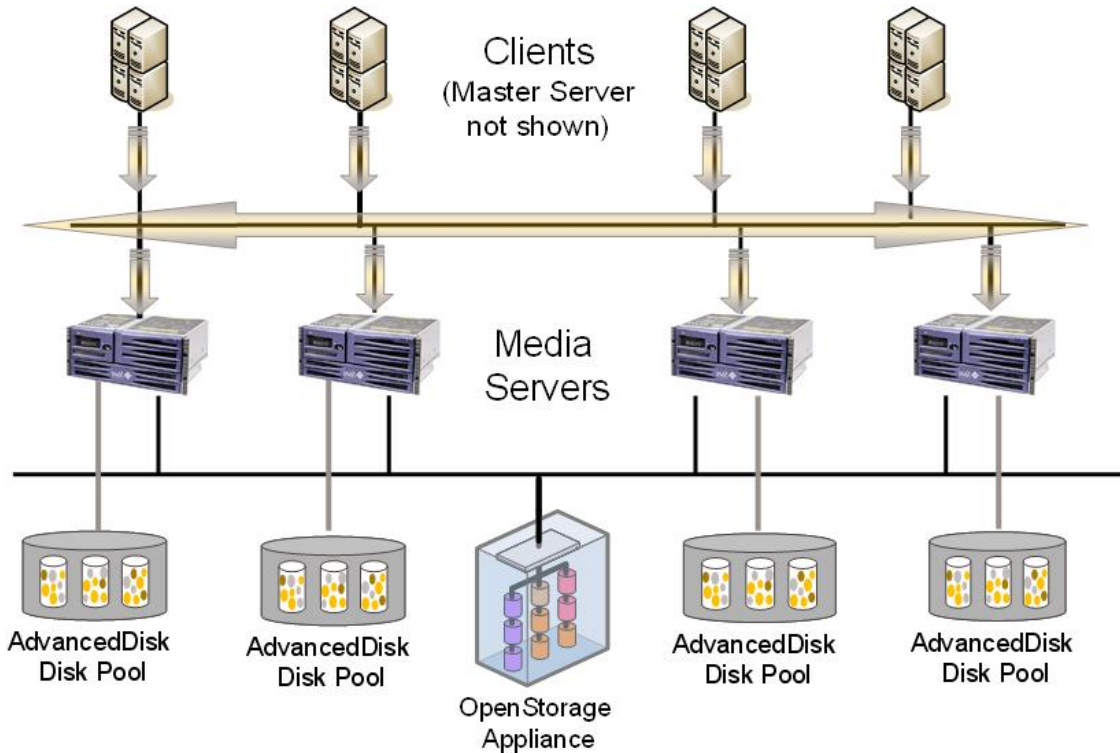


Figure 6 - Load balancing with AdvancedDisk and OpenStorage

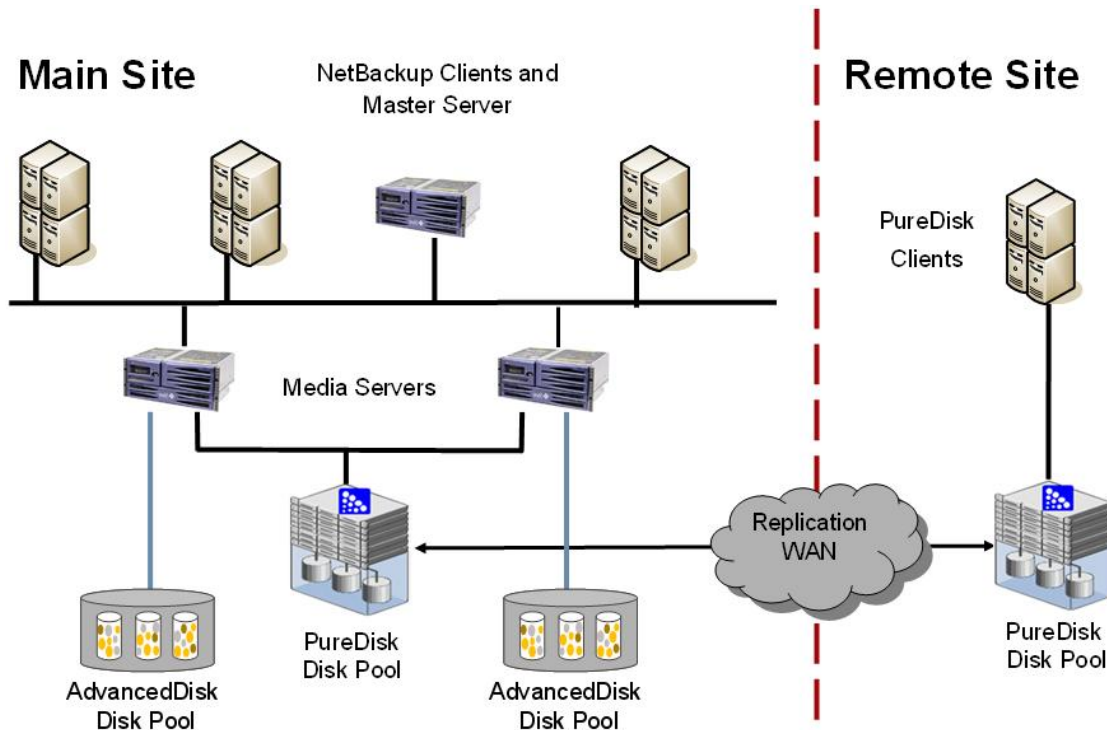
#### 4.2.2 Combined Main and Satellite Site Protection Using PureDisk

PureDisk Deduplication Option uses the same PureDisk infrastructure as the PureDisk Remote Office backup system. This means that a PDDO device located at a main data center can act as a remote replication server for a regular PureDisk server located at a satellite data center. **Error! Reference source not found.** shows a configuration of this nature. In the main data center backups from clients are directed via the Media Servers to AdvancedDisk Storage Units for short term storage. Using AdvancedDisk Storage Units with a Storage Unit Group ensures that the backups are load balanced across the available Media Servers. Backups that require longer term storage are duplicated to a PDDO device by Storage Lifecycle Policies.

At the satellite data center clients are backed up to a PureDisk storage server using PureDisk Remote Office. The PureDisk storage server at the satellite data center is replicated to the PDDO device at the main data center, providing off-site protection to the backup data. At the main data center these backups from the satellite location can, if required, be exported to NetBackup to allow rapid recovery of a failed client. This feature is particularly useful if the satellite site has no resident IT staff, as a failed client can simply be 'rebuilt' at the main site and shipped to the satellite site.

The backups from the main data center held on the PDDO device can also be replicated to the satellite data center to provide off-site protect. However as this occurs without NetBackup's knowledge backups replicated to the satellite site cannot be accessed directly by NetBackup and in the event of the PDDO device failing completely would have to be replicated back to the main data center before they can be used for restore.





**Figure 7 - Main and satellite site backup combination**

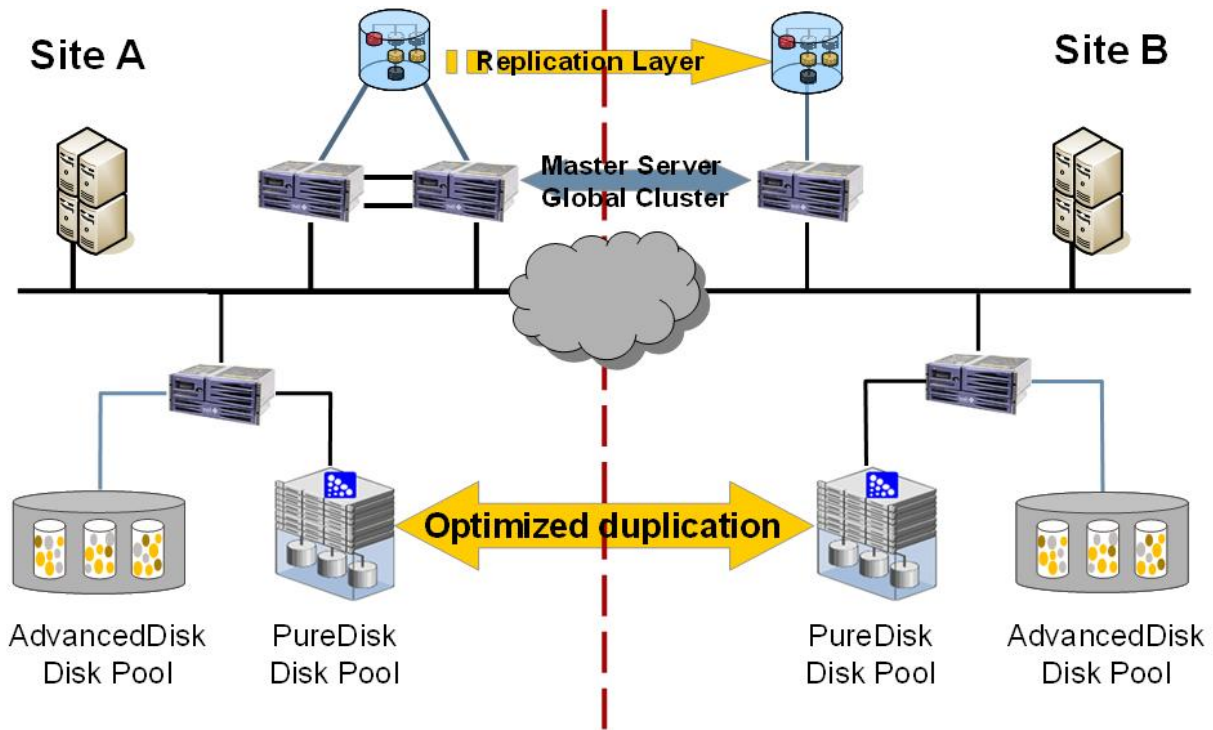
### 4.2.3 Highly Available Dual Site/Single Domain Configuration

The dual site/single domain model provides the ultimate in cost effective protection against site loss by combining two production sites into a single backup environment and ensuring that all critical backups exist in both locations. In a dual site/single domain configuration a single NetBackup domain spans two geographically remote sites, each of which effectively becomes the disaster recovery site for the other. Copies of the backups exist at both sites and the Master Server is clustered in such a way that control of the domain can be switched from site A to site B if necessary (this concept is discussed in more detail in a separate whitepaper, “Implementing Highly Available Data Protection with Veritas NetBackup”).

Figure 8 shows a Dual Site/Single Domain configuration with AdvancedDisk and PDDO storage. Backups are initially written to the AdvancedDisk storage at each site and the critical backups are then duplicated to the PDDO storage using Storage Lifecycle Policies. The PDDO devices are configured to replicate to each other and the optimized duplication feature is used to update the NetBackup catalogs to include the copy in the PDDO device at the remote site when the replication has completed.

In this configuration the PDDO storage is not used simply to provide long term space optimized storage, it is also used to provide an automatic ‘off site’ capability by leveraging the optimized duplication capability of PDDO.





**Figure 8 - Dual Site/Single Domain with PureDisk**

Under normal operation the Master Server runs on one of the cluster nodes on Site A. In the event of a site failure at Site A, the Master Server cluster is failed over to Site B. As the backups of the critical applications at site A have been duplicated to site B, recovery of those applications at site B can begin immediately. A full off-site disaster recovery of critical applications is achieved with no requirement to move any physical media between sites or to recover any NetBackup component before starting the application recovery.

## About Symantec

Symantec is a global leader in infrastructure software, enabling businesses and consumers to have confidence in a connected world. The company helps customers protect their infrastructure, information, and interactions by delivering software and services that address risks to security, availability, compliance, and performance.

Headquartered in Cupertino, Calif., Symantec has operations in 40 countries. More information is available at [www.symantec.com](http://www.symantec.com).

For specific country offices and contact numbers, please visit our Web site. For product information in the U.S., call toll-free 1 (800) 745 6054.

Symantec Corporation  
World Headquarters  
20330 Stevens Creek  
Boulevard  
Cupertino, CA 95014 USA  
+1 (408) 517 8000  
1 (800) 721 3934  
[www.symantec.com](http://www.symantec.com)

Copyright © 2007 Symantec Corporation. All rights reserved. Symantec and the Symantec logo are trademarks or registered trademarks of Symantec Corporation or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.

10/07 xxxxxxxx