



Data Protection Strategies for Network Appliance™ Storage Systems

Network Appliance, Inc.

August 2006 | TR-3066

Executive Summary or Abstract

Systems fail, users accidentally delete files, natural disasters occur, and mistakes happen. Throughout, businesses could lose critical data. One of the most important questions IT management needs to ask is, "What is my data recovery plan?" Network Appliance storage systems provide a variety of choices for data protection and recovery. This paper addresses the many available options and recommends solutions for protecting data on NetApp storage systems.

"Data Protection Strategies for Network Appliance Storage Systems" looks at data protection requirements, technology solutions, and performance issues and provides recommendations. The goal of this document is to help prepare for the future challenges of protecting against system downtime when the costs of unplanned outages are increasing along with requirements for continuous data access.

Disk storage capacity on servers is increasing at an alarming rate. Traditional data protection mechanisms are stretched to the limit. The lack of a backup window in many enterprises further escalates the problem. Storage capacity is scaling faster than tape cartridge capacity and tape bandwidth. Protecting multiterabyte storage systems with tape media can far exceed reasonable windows for both backup and restore. These trends suggest that additional methods complementary to tape backup should be considered.

Typically, businesses have five data protection requirements:

- Fast, user-initiated recovery of accidentally deleted files
- Tape archival of file systems or current unreferenced data for possible future use
- Minimized backup and recovery windows
- Fast recovery from natural or human-caused disasters
- Data protection for compliance regulation purposes and referencing

Network Appliance provides a unique set of solutions to address these requirements:

- Snapshot™ technology for daily online backups
- SnapRestore® software for near-instantaneous recovery of files or entire file systems to an earlier state
- SnapMirror® software for automated data replication
- [SnapVault®](#) software for disk-based backup and recovery of NetApp storage systems as well as open systems
- MetroCluster for disaster recovery over WAN
- Native dump and restore for backup and restore from tape
- Third-party data protection products with NDMP support for archiving data to tape
- SnapLock® and LockVault™ for compliance regulations
- VTL solution for faster backups by emulating tape devices on disk storage

Table of Contents

1. Introduction	5
2. Knowing Your Data	6
2.1 Business-Critical Data	6
2.2 Dynamic Data	6
2.3 Total Data Set Size	6
2.4 Number and Size of Files	7
2.5 Directory Structure	7
2.6 Data Types and Compression	7
3. Business Issues Affecting Data Protection	8
3.1 Identifying Business-Critical Data	8
3.2 Protecting Data from User Errors	8
3.3 Archiving Data for Future Use	8
3.4 Reducing Impact on Operations	8
3.5 Recovering from a Disaster	8
4. Data Protection Technologies for NetApp Storage Systems	9
4.1 Snapshot Technology for Online Backups	10
4.2 FlexVol™ Volumes for Data Protection	11
4.3 FlexClone Volumes in Data Protection	11
4.4 SnapRestore Software	11
4.5 SnapMirror Software: Automated Data Replication	12
4.6 MetroCluster: Site-Level Disaster Recovery	13
4.7 SnapVault Software: Disk-to-Disk Backup and Archiving	14
4.8 Native Backup Using Dump and Restore	15
4.9 Fibre Channel and Gigabit Ethernet Tape SAN Solutions	17
4.10 NearStore VTL	18
4.11 SnapLock for Compliance Data Protection	19
4.12 LockVault	20
4.13 Network Backup Using NFS Mounts and CIFS Shares	20
4.14 Remote Magnetic Tape (RMT) Protocol	21
5. Tape Data Protection Performance Recommendations	21

5.1 Concurrent Backup or Restore Sessions	21
5.2 Dealing with Different Data Types.....	22
5.3 CPU Load Concerns	22
5.4 Tape Drive Performance	22
5.5 Tape Devices.....	23
5.6 Tape Drive Recommendations.....	23
6. Summary of Recommendations.....	24
6.1 Choosing a Backup Software Technology	24
6.2 Configuring NearStore as a Backup Storage System.....	25
6.3 Organizing File System Data into Volumes and Quota Trees	27
6.4 Classifying Data by Value and Change Rate.....	27
7. Summary.....	27
Appendix	29

1. Introduction

Network Appliance enterprise storage systems reduce the cost and complexity of enterprise data management. NetApp storage systems running the Data ONTAP operating system consolidate data in a centralized location, eliminating the need for multiple general-purpose servers. NetApp unified storage systems provide simultaneous access to UNIX®, Windows®, and Web clients over TCP/IP and Fibre Channel networks. Built-in RAID-DP™ protects against data loss from disk failure. RAID-DP is an advanced, cost-effective disk failure/error protection solution protecting information in the event of a double disk outage within a single RAID group with no discernable performance impact and offers definite benefits over RAID1/0 and RAID6. In addition, clustered failover, SyncMirror® software, and redundant components further increase reliability.

No matter how reliable storage systems are, circumstances can conspire to result in loss of data, an organization's most valuable asset. Users can accidentally delete files, hardware failures can occur, and natural or human-caused disasters can bring down a data center. Strategic Research Corporation (SRC) surveyed 237 UNIX and PC sites for its 2000 Backup and Archive Profile report. Administrators at these sites performed an annual average of 263 single or multiple file restores and six full file system restores. Because data losses are a reality, planning for data recovery is imperative.

SRC also reports on a trend toward recentralization of servers, system management, and equipment purchasing, forced by the business need to keep networks continuously online, thus keeping data accessible to users. The SRC report states that the more centralized a network is, the lower the data management costs are. NetApp storage systems address this trend by providing reliable, consolidated file storage. Recentralization also presents opportunities for centralizing and automating data protection through the use of Network Appliance SnapVault software and third-party, centrally administered data protection solutions to further reduce cost and risk.

The traditional data protection mechanism is backing up data to tape. Projected technology trends suggest that additional methods complementary to tape backup should be considered. These trends are:

- The amount of data to be backed up is ever increasing
- Backup windows for many companies are shrinking or disappearing
- Storage capacity (number of disks in a system multiplied by the size of each disk) is increasing
- Storage capacity is scaling faster than both tape cartridge capacity and tape bandwidth

One of the real problems today is that storage capacity is scaling faster than tape cartridge capacity or tape bandwidth. As storage capacity increases, tapes become more and more impractical. Backing up a 6TB file system in compressed format requires approximately eight LTO Ultrium 3-tape drives. Restoring this same file system in a disaster recovery situation would take approximately 25 hours using a single LTO Ultrium 3-tape drive. Alternatively, if multiple tape drives running concurrent restores were added to reduce the restore window to a reasonable eight hours, you would need four tape drives. As file systems and storage systems in general grow in storage capacity, this problem becomes critical. While tape backup will continue to be an important part of most backup strategies, alternatives such as combining disk-to-disk backup or file system mirroring with tape backup must be considered.

Performance issues also affect data protection decisions. The section "Tape Data Protection Performance Recommendations" provides some suggestions on controlling factors that affect transfer rates for the technologies previously discussed. Finally, this document summarizes the recommendations on how to use NetApp storage systems to solve your data protection needs, including how to configure NearStore® as a backup storage system.

2. Knowing Your Data

Several key factors drive data protection strategies:

- Acceptable recovery windows for business-critical data
- How up to date restored dynamic data needs to be
- Total data set size and size of volumes and qtrees
- Number and size of files
- Directory structure
- Data types and compression

2.1 Business-Critical Data

One of the essential steps in determining a company's data protection strategy is to identify the relative priorities of the data in an enterprise and how quickly each type needs to be recovered in case of a disaster. One should also consider how up to date recovered files need to be in the event of a recovery operation. Categorizing data by how critical it is to the business allows system administrators to design flexible data protection strategies around restoration requirements.

Network Appliance recommends that customers optimize data protection by organizing file systems into multiple volumes and qtrees. For example, isolating critical data in its own volume or qtree allows them to:

- Mirror business-critical data so it is available for immediate disaster recovery
- Create frequent Snapshot copies so that recent versions of files are available online
- Limit the size of volumes or qtrees so tape backups and restores take less time

Separating critical data for purposes of backup and restore can also increase overall backup and restore rates for all data. Backing up critical data separately can reduce the backup frequency for data that is less critical. To know more about protecting data by enhancing resiliency, please refer to the [Storage Best Practices and Resiliency Guide](#).

2.2 Dynamic Data

Though it is similar to business-critical data, rapidly changing data may warrant a different protection strategy. For example, dynamic data sets may require more frequent backups to capture frequent changes. Isolating dynamic data in its own volume or qtree allows a business to design a protection strategy specifically around it.

For example, incremental backups of rapidly changing data will take longer than backups of more static data because there will be more changed data to back up. If data such as this is in its own volume, frequent incremental backups can be run. Infrequently changing archival data can be placed in a separate volume that gets occasional incremental backups and full backups only at widely spaced intervals.

How much change occurs with a data set also affects the use of SnapMirror and SnapVault technology. IT managers can schedule incremental transfers more frequently to ensure mirrored data and disk-based backups are up-to-date.

2.3 Total Data Set Size

Knowing the total data set size and the size of each volume and qtree lets system administrators estimate tape backup and restore windows and SnapMirror or SnapVault data transfer times and then decide whether the estimated windows are adequate for the business.

Table 1 calculates tape backup and restore windows using average backup and restore rates for a single LTO tape drive.

	500GB VOLUME	1.4TB VOLUME
Backup time @ 60GB/hr	8.3 hours	23.3 hours
Restore time @ 40GB/hr	12.5 hours	35 hours

Table 1) Sample backup and restore windows.

If tape backup or restore times are unacceptable, use faster tape drives if possible. Or administrators may want to consider dividing large volumes into smaller volumes or qtrees. For example, if a 1.4TB volume is divided into four qtrees, each qtree can be backed up to a separate tape drive, or separate full backups can be performed on four different nights.

For large volumes with long restore windows, consider using SnapMirror or SnapVault software for fast disaster recovery. Volume sizes over 1.4TB or total data set sizes greater than 4TB may exceed the natural performance limitations of SCSI or Fibre Channel and tape and may therefore be good candidates for a SnapMirror or SnapVault solution.

2.4 Number and Size of Files

For a given volume size, large numbers of small files take longer to process than small numbers of large files. More files mean a bigger directory structure to be processed. In addition, there is a fixed amount of overhead associated with each file being backed up, regardless of the size of the file. Knowing the number and size of files will give administrators insight into the tape backup or restore performance they can expect.

2.5 Directory Structure

Directory structure also affects the performance of a backup and recovery solution. Both large directories and deep directories decrease performance. Large directories increase the complexity of memory management when each directory is processed. The entire contents of a small directory can be loaded into memory at once, whereas only parts of a large directory can be in memory at one time due to memory size restrictions. Memory management and frequent reading of the directory from disk introduce overhead.

Deep directory structure generally means higher file-to-directory ratio. To contain a certain number of files, more directories are needed in a deep directory structure than in a flat one. More directories translate to more work when backing up the files. Moreover, at restore time, a child file or directory cannot be created unless its parent directory has already been created. This child-parent dependency prevents file and directory creation from being highly parallelized.

2.6 Data Types and Compression

Modern tape drives have built-in data compression mechanisms that attempt to compress the data stream, thereby speeding up the rate at which data is processed by the device. More compressible data yields faster backup rates. Less compressible data yields slower rates. Compression also compacts the data so that a tape holds more information.

Different data types can be compressed by different amounts. Text, such as newsgroup data, tends to have lots of redundancy and therefore can often be compressed as much as 1.5:1. In one backup performance test, transfer rates for highly compressible data (1.8:1) were as high as 98GB per hour.

The opposite extreme is noncompressible data such as graphics files or binary executables. Graphic objects tend to be compressed when they are created and cannot be further compressed. Highly compressed data

may actually expand slightly when written to tape. In the middle is mixed data, such as you might find in home directories that contain a mix of text, graphics, and binary files.

With all compression algorithms, attempts to further compress very dense data can result in slower backup rates than if the compression were turned off. Administrators may want to isolate dense data in its own qtree or volume and turn off compression in the drive for that particular qtree or volume.

3. Business Issues Affecting Data Protection

In planning data protection and disaster recovery strategies, organizations should first analyze the following business issues that help define requirements.

3.1 Identifying Business-Critical Data

The first step in creating a data protection plan is to identify business-critical information that needs to be recovered first in a disaster recovery scenario. Next, determine the relative priorities of the remaining data. The value of information should dictate the technologies used to protect it. In addition to the "business criticality" of information, factors such as data set size, file system structure, data type, and rate of change come into play.

3.2 Protecting Data from User Errors

Most businesses need the ability to retrieve accidentally deleted files quickly and with little effort. Files need to be backed up frequently because the recovered file should be as close as possible to the version that was lost. Ideally, end users should be able to perform online recovery operations by copying an earlier version of an accidentally deleted file from disk to their home directory. This capability frees the system administrator from having to restore a file from tape.

3.3 Archiving Data for Future Use

Archived data provides a complete, self-consistent replica of a collection of data, such as a project directory hierarchy, suitable for bringing back online at some future date. Many businesses require that archived tapes be stored at an off-site location, from which a file or file system can be recovered many years later. For example, a software development firm may need to reconstruct an early version of its software product to fix an error for a customer. This requires an archive-to-tape scheme from which earlier systems can be rebuilt.

3.4 Reducing Impact on Operations

Next, one should understand the backup window for your business. Do business operations allow for a backup window in which system usage is low and backups will not affect users? If so, when and how long is this backup window? If not, what is the acceptable impact of backup procedures on day-to-day operations? While restores tend to be infrequent, backups can be constant and should not appreciably degrade system performance.

3.5 Recovering from a Disaster

What are the costs per minute, hour, or day of a business shutdown resulting from a disaster? How long can the company afford not to be actively doing business? Realistically calculate the costs of lost user productivity, missed business opportunity, and system administrator data management time.

What is the restore window, or how fast must disaster recovery be? Will restoring from tape be possible within the restore window? In certain cases, disk-to-disk backup and recovery solutions may be sufficient to meet short restore windows. As restore windows continue to get smaller, some companies may require a server-mirroring solution wherein a system can be brought online much more quickly than previously possible with a single system. Still others might require a geographically separate, remote hot site that is available to go online within minutes of a disaster.

The next section presents data protection technologies available from Network Appliance. Armed with requirements and information on available technologies, businesses can begin to design and implement a data protection solution to fit their needs.

4. Data Protection Technologies for NetApp Storage Systems

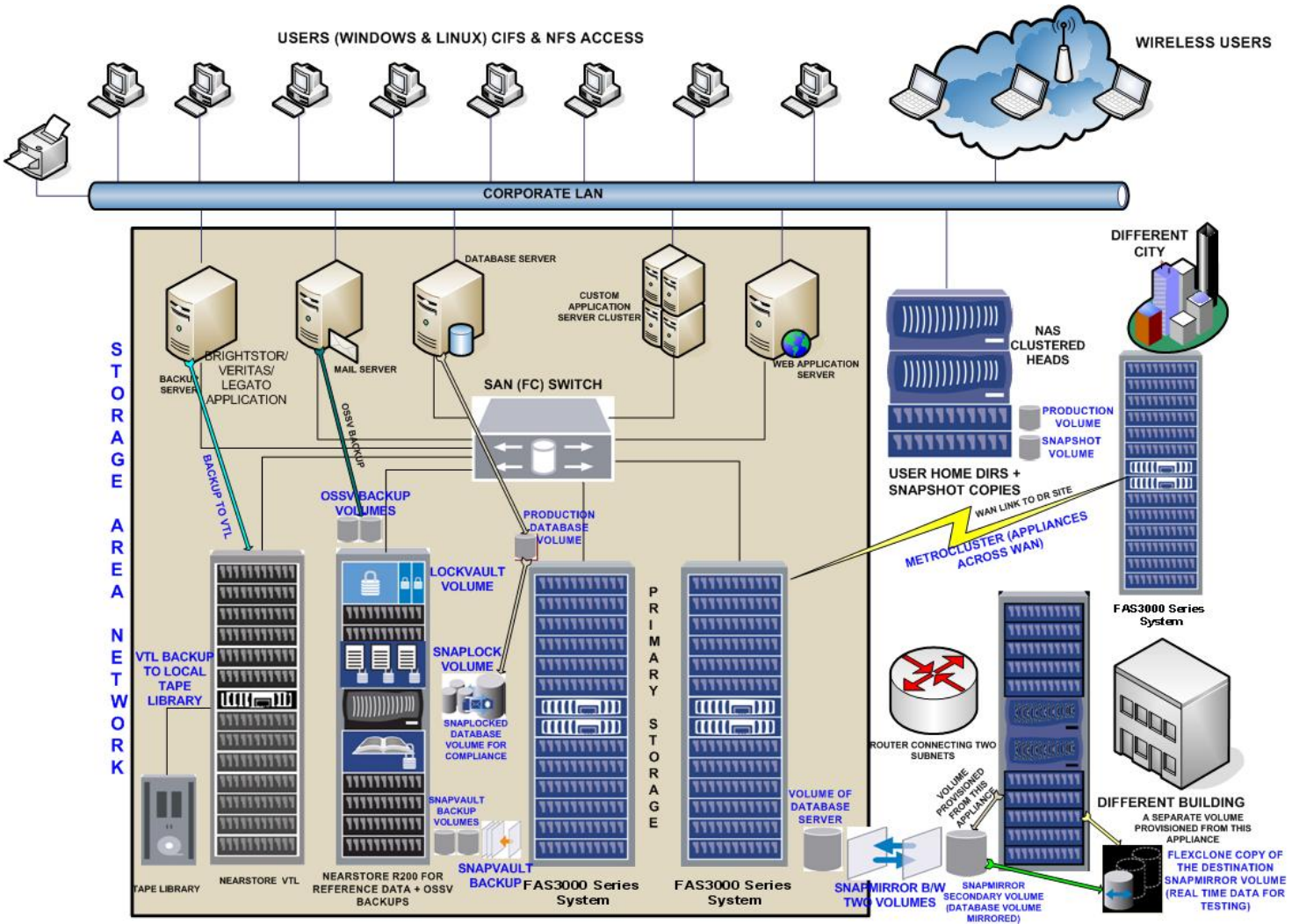


Figure 1) Data Protection Overview.

A high-performance and reliable data protection plan is critical to effective information technology operations. The Network Appliance WAFL® (Write Anywhere File Layout) file system provides unique data protection mechanisms, including, but not limited to, Snapshot, SnapRestore, SnapVault, and SnapMirror technologies. The flexible Network Data Management Protocol (NDMP) is an open, standard protocol enabling data protection vendors to incorporate Network Appliance storage systems with their full-featured solutions into centralized, distributed, or hybrid data protection models. Network Appliance and third-party technologies allow for a broad range of backup strategies that can satisfy virtually all enterprise-wide backup

requirements. A graphic of the role played by various data protection solutions from NetApp, which will be discussed in this section, is shown in Figure 1

4.1 Snapshot Technology for Online Backups

The Network Appliance WAFL file system supports Snapshot, a unique feature that allows administrators to maintain read-only versions of each file system online. Snapshot copies, a bundled component of Data ONTAP® software, allow users to recover accidentally damaged or deleted data by copying a desired file from a Snapshot directory. Versions of Data ONTAP 6.4 or greater currently support 255 Snapshot copies per volume.

Snapshot copies augment and simplify an overall enterprise data protection strategy. Snapshot copies can serve as daily backups from which users can recover their own files. The system administrator defines a schedule of hourly, nightly, or weekly Snapshot copies and defines how long to retain each Snapshot copy. By creating Snapshot copies throughout the day (every three hours from 8 a.m. to 8 p.m., for example) a system administrator can guarantee that recent file versions are available for recovery. A business may also choose to create nightly Snapshot copies and retain them for one week instead of nightly incremental backups to tape.

Tape backup, SnapRestore, SnapMirror, and SnapVault, all described later in this document, also use the Snapshot feature. The dump command and NDMP-compliant products read backup data directly from a Snapshot copy, eliminating the need to take the file system offline or deal with open file conflicts. SnapRestore software allows the administrator to nearly instantaneously revert a file or an entire file system. The SnapMirror automated replication software uses Snapshot copies to provide asynchronous mirroring. SnapVault software uses Snapshot copies to provide disk-to-disk block-level incremental backup and archive capabilities to Network Appliance storage systems.

One might assume that Snapshot copies would incur a significant disk space penalty because each Snapshot copy appears as though it is a read-only copy of the file system. However, in reality, Snapshot copies usually only require a small disk space premium. Snapshot copies are maintained as pointers to disk blocks containing data. When the WAFL file system creates a Snapshot copy, it makes a copy of the set of pointers from the active file system, but does not copy data blocks. As the active file system changes, Snapshot copies continue to point to deleted or changed disk blocks, holding these blocks from the file system's free space, thereby using disk space.

There are cases in which a business may choose to perform daily tape or SnapVault backups as well as create daily Snapshot copies. For example, in the unlikely occurrence of a concurrent failure of three disks in a single RAID-DP RAID group, data may not be recoverable from online Snapshot copies. For very valuable data or longer-term storage, a business may decide to perform nightly tape backups or SnapVault updates.

In summary, Snapshot copies provide:

- User-initiated recovery of accidentally deleted files
- Replacement for nightly incremental backups to tape
- Ability to save data more frequently than incremental backups to tape
- A consistent copy of the file system for dump and NDMP to use when creating tape backups
- Underlying technology for SnapRestore, SnapMirror, and SnapVault software, as described in the following sections

4.2 FlexVol™ Volumes for Data Protection

With the introduction of FlexVol technology, data protection has gained significant advantage. Data can now be placed on FlexVol volumes for a good combination of performance, effective usage of disks, and data protection. Here is a brief explanation of how it helps.

FlexVol volumes are created on top of a large pool of disks called an aggregate. There can be more than one aggregate if required. FlexVol volumes are striped across every disk in the aggregate and have their own attributes, which are independent of each other. For example, they can have their own Snapshot schedule or their own replication schedule.

Also, FlexVol volumes can also be increased or decreased in size on-the-fly. They also have another very important attribute. Space that is allocated to FlexVol but not used can be taken away, on-the-fly, and reallocated to another FlexVol volume that needs it. The aggregate(s) can also be increased in size on-the-fly. FlexVol volumes can also be cloned using FlexClone™ technology. A FlexClone volume represents a space efficient point-in-time copy (read/write) of the parent FlexVol volume but can also be turned into a fully independent FlexVol volume itself.

For more comprehensive information on FlexVol volumes, please refer to Technical Report 3356 at www.netapp.com/library/tr/3356.pdf.

4.3 FlexClone Volumes in Data Protection

FlexClone volumes have all the capabilities of a regular flexible volume, including growing, shrinking, and being the source for Snapshot copies or even the source for another clone.

FlexClone volumes also enable administrators to access the destination mirror created through the NetApp SnapMirror product. Previously, it was necessary to break the mirror in order to make any changes to the destination copy. With FlexClone, an administrator can now clone a Snapshot copy held in the mirror and make it available for both reading and writing at the remote site while allowing the mirror facility to continue running unaffected.

For more detailed information on FlexClone volumes please refer to Technical Report 3347 at www.netapp.com/library/tr/3347.pdf.

4.4 SnapRestore Software

The SnapRestore software leverages the Snapshot feature of Data ONTAP software by restoring a file or entire file system to an earlier preserved state. It can be used to recover a damaged or deleted file or to recover from a corrupted database, application, or damaged file system.

When running Data ONTAP 6.2 or above, the system administrator can restore a file or the entire file system from any existing Snapshot copy. Without rebooting, the restored file or volume is available for full production use, having returned to the precise state that existed when the selected Snapshot copy was created. SnapRestore software can be used in conjunction with native restore commands when you need to perform both a full tape restore and one or more incremental restore operations. Since restoring from tape sometimes fails, the following process can be used to increase reliability:

1. Perform full restore
2. Create Snapshot copy A
3. Perform incremental restore

Note: If step 3 fails, use SnapRestore software to revert to Snapshot copy A and start step 3 again.

Without SnapRestore software, you would need to destroy and then recreate the volume and start again with step 1.

4.5 SnapMirror Software: Automated Data Replication

SnapMirror software provides a fast, flexible enterprise solution for mirroring or replicating data over local or wide area networks. SnapMirror can be used for:

- Disaster recovery
- Remote enterprise-wide online backup
- Data replication for local read-only access at a remote site
- Application testing on a dedicated read-only mirror
- Data migration between Network Appliance storage systems

SnapMirror technology is a key component of enterprise data protection strategies. If a disaster occurs at a source site, businesses can access mission-critical data from a mirror on another NetApp storage system, ensuring uninterrupted operation. Enterprise tape backups can be made from the mirror rather than a production system, reducing CPU load on the production system.

The destination NetApp storage system can be located virtually any distance from the source. It can be in the same building or on the other side of the world, as long as the interconnecting network has the necessary bandwidth to carry the replication traffic that is generated.

SnapMirror software provides very fast recovery in a disaster situation compared with restoring a file system or quota tree (qtree) from tape. To assist customers in determining the economic impact to their company that would result from total system downtime, Network Appliance recommends that customers realistically calculate costs per day of business shutdown resulting from a disaster. As the cost of downtime rises for an organization, enterprises cannot afford to operate without a disaster recovery solution in place. SnapMirror allows organizations to quickly and easily implement an economical and reliable disaster recovery solution.

SnapMirror technology leverages the WAFL Snapshot capability to create and update a copy of a source volume or qtree on a destination NetApp storage system. The mirror copy is accessible to users in read-only mode on the destination NetApp storage system. SnapMirror software makes a baseline transfer of the data (comparable to a full backup for tape backups). The initial transfer can either be accomplished through a network connection or through the restore of a tape on the destination. SnapMirror then updates the mirror by replicating only new or changed data blocks. Mirror copies are consistent because SnapMirror software operates on consistent Snapshot copies.

System administrators specify the intervals at which SnapMirror Snapshot copies are created and the times at which incremental transfers will occur. Determining this schedule depends upon how much the data changes during the day, how up-to-date the mirror needs to be, CPU usage on the source, and available network bandwidth.

SnapMirror can operate in three different types of modes: asynchronous, synchronous, and semi-synchronous. In asynchronous mode, SnapMirror performs incremental, block-based replication as frequently as once per minute. Performance impact on the source FAS system is minimal, as long as the system is configured with sufficient CPU and disk I/O resources. Because asynchronous replication is periodic, SnapMirror is able to consolidate writes and conserve network bandwidth. There is minimal impact on write throughput and write latency. As soon as data is written to the NVRAM of the source system, applications using this data are free to continue processing, without waiting for the data to reach a destination system. Updates take place in the background, so the application does not experience any additional transaction latency.

Certain environments have very strict uptime requirements. All data that is written to one site must be mirrored to a remote site or system synchronously. SnapMirror in synchronous mode is a mode of replication that sends updates from the source to the destination as they occur, rather than according to a

predetermined schedule. This guarantees that data written on the source system is protected on the destination even if the entire source system fails. In synchronous mode, SnapMirror immediately replicates all data written to the source file system. This guarantees zero data loss in the event of a failure, but can have a significant performance impact. It is not necessary or appropriate for all applications.

A semi-synchronous mode is also provided which minimizes data loss in a disaster while also minimizing the extent to which replication impacts the performance of the source system. Unlike asynchronous mode, which can replicate either volumes or quota trees, synchronous and semi-synchronous modes work only with volumes. Complete details can be found in the Synchronous SnapMirror Design and Implementation Guide (TR-3326), available on the Network Appliance corporate Web site.

In summary, SnapMirror software can be used for:

- Data replication to a local or remote site
- Fast recovery from disaster; no lengthy restores from tape are required
- Replicating data to remote systems where it can then be backed up to tape
- Creating writable, FlexClone copies of real-time data, useful in application testing and development
- Migrating data between NetApp storage systems

4.6 MetroCluster: Site-Level Disaster Recovery

MetroCluster software provides an enterprise solution for high availability over wide area networks.

MetroCluster deployments of NetApp storage systems are used for:

- Business continuance
- Disaster recovery
- Achieving recovery point and/or recovery time objectives

MetroCluster technology is an important component of enterprise data protection strategies. If a disaster occurs at a source site, businesses can continue to run and access data from a clustered node in a remote site.

The primary goal of MetroCluster is to provide mission-critical applications redundant storage services in case of site-specific disasters. It is designed to tolerate site-specific disasters with minimal interruption to mission-critical applications and zero data loss by synchronously mirroring data between two sites.

A MetroCluster system is made up of the following components and requires the following licenses:

- **Multiple storage controllers, HA configuration.** Provides automatic failover capability between sites in case of hardware failures
- **SyncMirror.** Provides an up-to-date copy of data at the remote site; data is ready for access after failover without administrator intervention
- **Cluster remote.** Provides a mechanism for administrator to declare site disaster and initiate a site failover via a single command for ease of use
- **FC switches.** Provide storage system connectivity between sites that are greater than 500* meters apart

MetroCluster allows the active/active configuration to be spread across data centers up to 100 kilometers apart. In the event of an outage at one data center, the second data center can assume all affected storage operations lost with the original data center. SyncMirror is required as part of MetroCluster to ensure an identical copy of the data exists in the second data center should the original data center be lost.

1. MetroCluster along with SyncMirror extends active/active Clustering across data centers up to 100 kilometers apart
2. MetroCluster and SyncMirror provide the highest level of storage resiliency across a local region
3. Highest levels of regional storage resiliency ensure continuous data availability in a particular geography
4. NetApp recommends MetroCluster be used with SyncMirror for the highest level of storage resiliency

4.7 SnapVault Software: Disk-to-Disk Backup and Archiving

SnapVault is a data protection solution for heterogeneous storage environments, providing online disk-to-disk backup and recovery for NetApp storage systems as well as open systems. SnapVault provides intelligent data movement, reducing network traffic and impact on production systems, performing frequent backups to ensure superior data protection using field-proven NetApp Snapshot technology to efficiently store days' and potentially weeks' worth of backups online. Snapshot copies can then be backed up to tape using any NDMP-compliant backup software application.

A full backup copies the entire data set to backup media, which is tape in traditional backup applications or a NearStore system when using SnapVault. An incremental backup copies only the changes in a data set to backup media. Because incremental backups take less time and consume less network bandwidth and backup media, they are less expensive. Of course, since an incremental backup contains only the changes to a data set, at least one full backup is required in order for an incremental backup to be useful.

SnapVault software provides the following features:

- Dramatically reduced backup windows versus traditional tape-based backup
- Guaranteed integrity for backup data, eliminating recovery failures due to media errors
- Remote enterprise-wide online backup
- User-initiated recovery of accidentally deleted files
- Fast recovery of corrupted or destroyed data
- Reduced tape media utilization from decreased reliance on tape

SnapVault software, like SnapMirror, leverages Snapshot copies to transfer data from Network Appliance storage systems or open systems servers, referred to as SnapVault primaries. The data is transferred to another Network Appliance storage system, typically a NearStore appliance, which is referred to as the SnapVault secondary. The data on the secondary is accessible to users in read-only mode. SnapVault software makes a baseline transfer of the data (comparable to a full backup for tape backups). When a NetApp storage system is the primary system, incremental updates to data will consist of new or changed data blocks. When an open storage system is the primary, incremental updates to data will consist of only new or changed files.

One of the unique benefits of SnapVault is that users do not require special software or privileges to perform a restore of their own data. Any users who wish to perform a restore of their own data may do so without the intervention of a system administrator, saving time and money.

Restoring a file from a SnapVault backup is simple. Just as the original file was accessed via an NFS mount or CIFS share, the SnapVault secondary may be configured with NFS exports and CIFS shares. As long as the destination qtrees are accessible to the users, restoring data from the SnapVault secondary is as simple as copying from a local Snapshot copy.

System administrators specify the intervals at which SnapVault Snapshot copies are created on NetApp SnapVault primary systems and the times at which incremental transfers will occur. Updates can also be

initiated manually via a command-line interface. Additionally, the business continuance module for Network Appliance DataFabric® Manager can be used to create and modify these schedules.

SnapVault stores backup copies of the data in the WAFL file system, which replicates all of the file permissions and access control lists held by the original data; if a user was not authorized to access a file on the original file system, that user will not be authorized to access the backup copies of that file. This allows the self-service restores described above to be performed safely.

Administrators can utilize SnapVault to transfer qtrees from NetApp storage systems and directories from open systems storage to NearStore and create tape backups from NearStore. Administrators can also centralize tape backup of geographically dispersed sites by transferring data to one central location and backing up all data sets to tape from a single location, reducing the resources required at each remote site to support local data protection.

SnapVault has been extended to support heterogeneous systems via Open Systems SnapVault. With OSSV, servers running Windows, Sun™, UNIX, and Linux® operating systems can be backed up to a NetApp storage appliance directly, with these servers configured as OSSV primary backup sources.

4.8 Native Backup Using Dump and Restore

The Network Appliance native dump and restore commands form the foundation of a scalable tape backup strategy. Dump and restore, also bundled components of Data ONTAP software, are optimized in the Data ONTAP kernel for efficiency and reliability. These commands can be used in simple console-based backups of small volumes copied to a single high-capacity tape. Or they can be invoked from an NDMP-based product for complex enterprise-wide backup of multiple volumes on multiple NetApp storage systems.

Dump writes file system data to backup media in an archival format that can be restored on NetApp storage systems by native restore. Since the Data ONTAP dump command uses a format that is compatible with Berkeley BSD-based dump commands, these archives can also be restored on a Sun Solaris™ system by `ufsrestore`. Dump saves a consistent view of the file system by copying data from a Snapshot copy. Dump can back up a full volume, a qtree, any specific directory, or an individual file. Restore can restore a full volume, a qtree, a directory, or individual files. Dump and restore will back up and recover NFS and CIFS (or UNIX and Windows NT®) file attributes. Restoring a Network Appliance dump tape results in a consistent file system or subset of a file system, including NFS and CIFS file attributes.

In summary, native dump and restore provide:

- Full file system, subdirectory, and single file restores
- Data recovery on Sun Solaris with `ufsrestore`
- The ability to make data archives for off-site storage
- Backup to and recovery from locally attached tape devices for high performance

Third-Party NDMP-Based Data Protection Solutions

Network Data Management Protocol (NDMP) (www.ndmp.org) is an open standard for centralized control of enterprise-wide data management. NDMP enables backup software vendors to provide support for Network Appliance storage systems without having to port client code. An NDMP-compliant solution separates the flow of backup and restore control information from the flow of data to and from backup media. These solutions invoke Data ONTAP software's native dump and restore to back up data from and restore data to a NetApp storage system.

NDMP-based solutions can centrally manage and control backup and recovery of highly distributed data while minimizing network traffic. These products can direct a NetApp storage system to back itself up to a locally attached tape drive, without sending the backup data over the network. NDMP-based solutions are designed to assure data protection and efficient restoration in the event of data loss and include many

control and management features not available with a NetApp storage system's native dump and restore commands, such as discovery, configuration, scheduling, media management, tape library control, and user interface.

Network Appliance partnered with Intelliguard, part of Legato (now EMC), to create NDMP. The two companies have promoted its standardization in the industry. Some examples of industry leading backup software vendors who support NDMP and Network Appliance storage devices are listed here:

- Atempo® Time Navigator (www.atempo.com)
- BakBone® NetVault® (www.bakbone.com)
- CommVault® Galaxy (www.commvault.com)
- Computer Associates™ BrightStor™ ARCserve Backup (www.ca.com)
- HP OpenView Storage Data Protector (www.hp.com)
- IBM Tivoli Storage Manager (www.ibm.com)
- EMC Legato® NetWorker® (www.legato.com)
- Syncsort® Backup Express (www.syncsort.com)
- VERITAS® NetBackup™ (www.symantec.com)

The NDMP specification allows the following backup configurations:

- Local backup from a NetApp storage system to a direct-attached tape device
- Three-way backup from a NetApp storage system via the network to another NetApp storage system with a local tape device
- Backup from a UNIX or Windows NT server via the network to a NetApp storage system with a local tape device
- Backup from a NetApp storage system via the network to a UNIX or Windows NT backup server with a local tape device

Table 2 specifies the backup configurations each partner supports.

	LOCAL FAS TO TAPE	FAS SYSTEM TO FAS SYSTEM	SERVER TO FAS SYSTEM	FAS SYSTEM TO SERVER
Atempo Time Navigator	X	X	X	X
BakBone NetVault	X	X	X	X
CommVault Galaxy	X	X		X
Computer Associates BrightStor ARCserve Backup	X	X		
HP OpenView Storage Data Protector	X			
IBM Tivoli Storage Manager	X			
EMC NetWorker	X	X		X *
Syncsort Backup Express	X	X		X

Symantec® NetBackup	X	X		X
* Available with purchase of SnapImage v1.5 with NetWorker 6.0.1.				

Table 2) Capabilities of NDMP-compliant backup partners.

NDMP-based third-party solutions provide:

- Central management and control of highly distributed data
- Local backup of NetApp storage systems without sending data over the network
- Control of robotics in tape libraries
- Data protection in a mixed server environment of UNIX, Windows NT, and NetApp storage systems
- Investment protection in established backup strategies

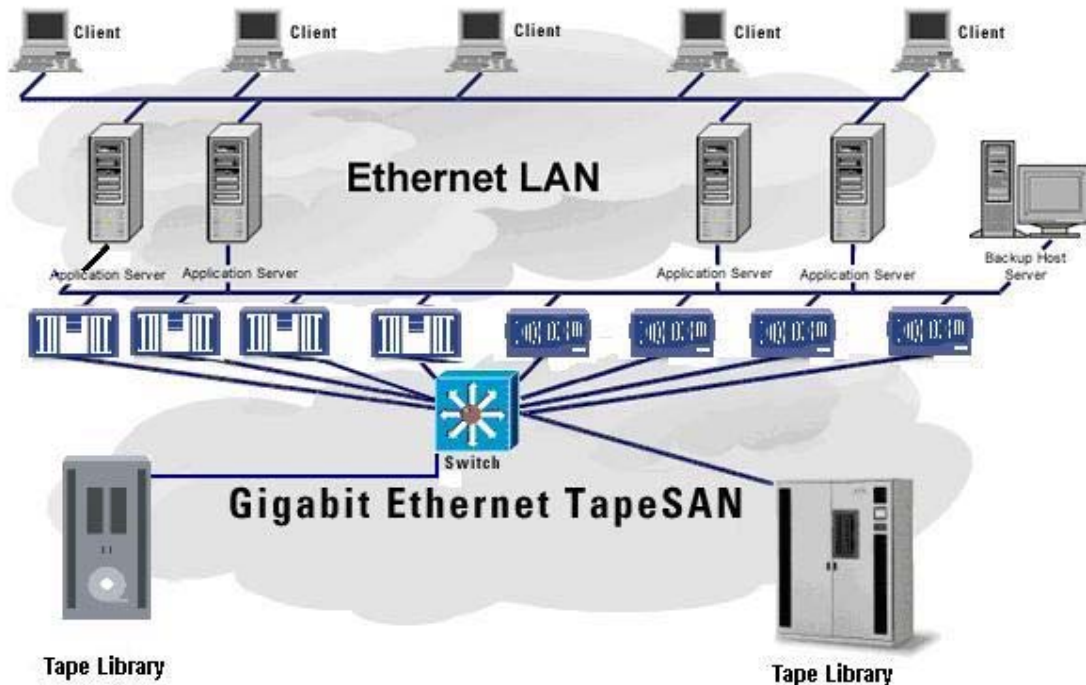
4.9 Fibre Channel and Gigabit Ethernet Tape SAN Solutions

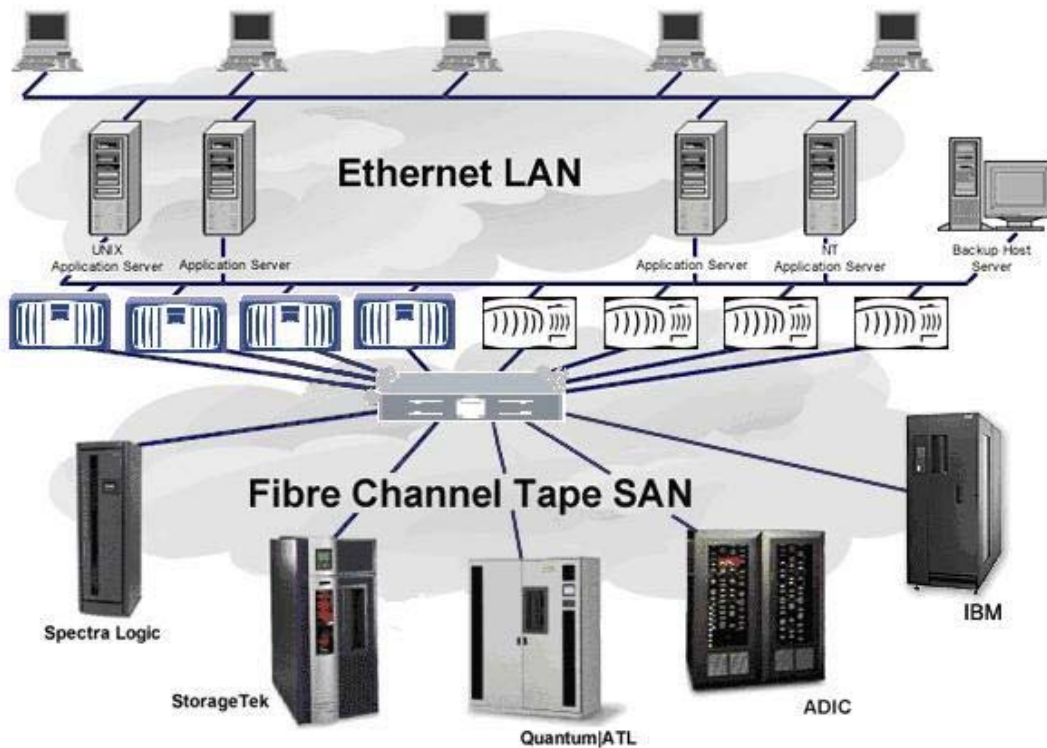
Network Appliance delivers both certified Fibre Channel Fabric Tape SAN backup solutions and Gigabit Ethernet Tape SAN solutions. These solutions are made possible through NetApp joint partnerships with industry leaders in the fields of tape automation, fabric switches, and backup software. They offer significant benefits for enterprise customers over tape devices attached directly to NetApp storage systems via SCSI. Specifically, both solutions offer the following benefits:

- Tape sharing and amortization of tape resources
- Extended distances from data to centralized tape backup libraries
- Minimized impact from backups on servers on the network
- Tape drive hot-swapping
- Dynamic tape configuration changes without shutting down the NetApp storage system

Figure 2 illustrates the Gigabit Ethernet Tape SAN configuration.

Figure 2) Gigabit Ethernet Tape SAN.





Gigabit Ethernet Tape SAN configurations allow multiple Network Appliance storage systems to concurrently transfer data over Gigabit Ethernet to one or more tape libraries that support NDMP. This architecture allows each drive inside the tape library to be seen as a shared resource and as an NDMP server. One clear advantage of this configuration is the demonstrated interoperability of Ethernet-based components.

Figure 3) Fibre Channel Tape SAN.

Together with a third-party NDMP-based data protection solution that supports technology known as dynamic drive sharing, both Fibre Channel and Gigabit Ethernet Tape SAN solutions enable customers to dynamically allocate tape drives in a larger library to NetApp storage systems as needed for backup or recovery operations. This eliminates the need to dedicate expensive tape devices to each system.

These solutions both help provide essential elements to enterprise customers seeking to maximize the availability of their Network Appliance storage. Customers can replace or upgrade tape devices with no impact on the system's ability to serve data to clients. Drives can be dynamically added or even removed without requiring any downtime. For more specific information about these solutions, as well as a list of currently certified equipment, visit the Open Storage Networking Initiative.

4.10 NearStore VTL

In order to address the issues related to backup windows and performance, NetApp NearStore Virtual Tape Library (VTL) emerges to be a faster, more cost-effective, more reliable, and easier-to-manage replacement for tape libraries. The NearStore VTL device enables a disk storage system to appear exactly like one or

more tape libraries to the backup application. It is the easiest way to incorporate disks into a backup environment. This enables customers to rapidly deploy disks for their backup and restore needs and begin enjoying the benefits of disk speeds—faster recovery times and simplified management—immediately. At the same time, this VTL solution integrates with real physical tape libraries for the seamless creation of removable media sets.

With NearStore VTL, the user defines the number of virtual tape drives that can be used by the backup application like any other tape device. VTL offers the following advantages over traditional tape backup:

- Elimination of half-written tapes for better space utilization by the NearStore VTL proprietary tape smart sizing technology
- Management of data on disk instead of tape
- Fast and easy expiration of tape cartridges
- Creation of extremely large virtual tapes to simplify backups of large catalogs that cannot span tapes
- Faster throughput than tape libraries (since this is disk based)
- NearStore VTL uses self-tuning technology for better disk utilization and better throughput; this reduces the work of system administrators as manual intervention is eliminated

Adding to this NearStore VTL can be incorporated with NetApp data protection technology using the Decru appliance that can encrypt data that are moved from server to NearStore VTL or from NearStore VTL to the physical tape library this enhances data security.

NearStore VTL offers a number of functionality benefits over backing up data directly to disk. It manages free space better than the backup applications today enabled with Tape Smart Sizing and self-tuning technology. All these do not require a change in the backup application configuration as it is tested and certified with all leading backup software vendors. VTL also enables greater connectivity to backup data. By using this type of backup strategy, instead of having local tape libraries at every site, IT administrators can eliminate the need for shipping tape media to their disaster recovery operation sites.

The NearStore VTL solution can be used as a primary backup stage before passing the data to tape and the tape library can be used as a secondary backup device,

As all NetApp FAS systems have visibility to all tape devices in the SAN, these backup solutions enable dynamic tape sharing and optimal utilization of tape resources through Data ONTAP software and through NDMP software applications that support dynamic tape sharing.

Note that VTL appliances cannot manage NDMP dump directly from NearStore to tape that is SAN-attached.

NearStore VTL backup solutions typically comprise an ATA/SATA disk array with NearStore head and NetApp proprietary VTL software that make the array to emulate tape. A VTL acts just like a tape library, providing an easy way to move from tape to disk-based backup without any process changes.

NearStore VTL solution has a programmable file system that allows users to lay out volumes on disk in a way that emulates tape. It also provides an easier and faster means of getting data from the VTL to the tape devices, as the data is typically already laid out on disk in the proper format for the tapes. Cloning or vaulting the data is faster than performing a second backup to tape.

4.11 SnapLock for Compliance Data Protection

NetApp SnapLock enables compliance with regulatory and best-practices records-retention requirements by allowing the creation of nonrewritable, nonerasable WORM volumes on NetApp NearStore and FAS storage systems, thereby preventing critical files from being altered or deleted until a specified retention date.

SnapLock allows WORM data to be replicated securely and automatically between multiple NetApp NearStore and FAS systems using NetApp SnapMirror software. WORM-to-WORM replication enables data at remote sites to fully comply with regulations or best practices, resulting in a highly robust WORM data protection solution. WORM data can also be backed up to tape for an additional level of data protection.

Open, Standards-Based Solution for Easy Integration

NetApp SnapLock supports open, industry-standard protocols such as NFS and CIFS, so it offers easy data access and application integration—as opposed to other regulated data solutions that require each application vendor to write to a closed, proprietary API to store, retrieve, and search data. Consolidating storage on unified, rapidly scalable NetApp NearStore or FAS systems enables increased flexibility in preserving important data using archival and records-management applications.

4.12 LockVault

LockVault integrates NetApp SnapLock and NetApp SnapVault technologies to create the only solution specifically designed to address regulatory compliance requirements for unstructured data. A NetApp LockVault solution eliminates the need to continuously identify and classify each file that might fall under regulatory jurisdiction. LockVault allows customers to make permanent WORM backups of their unstructured data, by simply copying and storing only the unique blocks that have been written since the most recent incremental backup. Every block variation is protected, so no data falls through the cracks. LockVault also automatically provides online compliant backups of unstructured data at multiple points in time. As a result, any backup can be searched, indexed, or instantaneously recovered. Each backup is immutable until its specified expiration date. In addition, NetApp SnapMirror fully supports WORM-to-WORM remote replication to easily provide the required duplicate copy of the compliant data. This technology is fully compatible with Windows, UNIX, or Linux data residing on any NetApp system

With a NetApp solution, one can make unalterable copies of all file servers on a regular basis as well as mirrored copies of all compliant data as often as desired. Unlike other compliant media (including tape), LockVault permits block-level incremental backup, provides rapid discovery and retrieval capabilities, combines compliance and data protection, and improves data management across the enterprise. NetApp LockVault technology creates compliant online backups of unstructured data that dramatically reduce the amount of storage required while optimizing data retrieval. Today's financial industry is facing significant IT challenges. NetApp provides a complete, easy-to-manage, cost-effective solution to meet all 17a-4 compliance requirements. The NetApp LockVault solution is easy to deploy and requires no change in business processes. With a NetApp solution, a company can address the unstructured data compliance problem immediately and fully, placing it on the fast track to regulatory compliance.

4.13 Network Backup Using NFS Mounts and CIFS Shares

Network backup involves mounting or mapping an export or share by a backup server that has a high-capacity tape drive or tape library directly attached. Using virtually any backup application, all files under the mounted/mapped export/share are subsequently copied over the network to the backup server, where they are immediately transferred to the attached tape device.

This backup method provides flexibility in choosing which enterprise-wide backup application to use. It allows virtually any backup application to back up data on Network Appliance storage over a network connection. However, this method can be significantly slower than backup to locally attached tape devices.

4.14 Remote Magnetic Tape (RMT) Protocol

The RMT protocol is an industry-standard protocol that enables you to write a backup data stream over a network connection from NetApp storage to a remote tape drive attached to a remote host. The remote host must support the RMT protocol. The RMT protocol is a bundled component of Data ONTAP software.

Remote backup via RMT allows easier sharing of one or more tape drives among a cluster of NetApp storage systems. In addition, the RMT protocol provides access to nonqualified tape drives. You can attach these tape drives to a remote host that supports both the drives and the RMT protocol. Nevertheless, RMT devices have more overhead and can be expected to be slower than locally attached devices.

5. Tape Data Protection Performance Recommendations

Network Appliance testing has determined transfer rates for dump and restore are based on the following factors in decreasing order of magnitude:

- Number of concurrent backups or restores
- Type of tape drive
- Data types and compression
- CPU load

NetApp testing showed that the following factors had little to no impact on dump and restore transfer rates on Network Appliance FAS systems:

- Disk drive capacity
- Volume size
- Disk count (or spindle count)

It should be noted that due to significant architecture differences in the NearStore product line, these three factors do have a significant impact on different models of the NearStore family. For performance recommendations specific to NearStore, review Best Practices Guide for Tape with NearStore Appliances.

Network Appliance has qualified a variety of tape drives and Fibre Channel SAN and Gigabit Ethernet SAN-attached tape libraries, as listed in the Data Protection Solutions page of the NetApp Web site at www.netapp.com/solutions/data_protection.html. The Network Appliance data protection partners have qualified additional tape libraries. Most of these partners publish tape-device qualification tables on their Web sites.

5.1 Concurrent Backup or Restore Sessions

The largest single factor in improving tape-based data protection performance is increasing the number of simultaneous backup or recovery sessions, especially when using locally attached tape drives. The maximum numbers of concurrent backup or restore sessions supported by various NetApp storage systems are listed in Table 3.

NETAPP STORAGE SYSTEM	CONCURRENT BACKUP/RESTORE SESSIONS
FAS900 Series	FAS920 – 8, FAS940 – 16, FAS960 - 16
FAS3050 Series	16
FAS3020 Series	8 (Data ONTAP 7.0.1), 16 (Data ONTAP 7.0.1.1 and later)

NearStore	128
FAS270	8

Table 3) Concurrent backup/restore sessions per NetApp storage system.

Note: The actual numbers of concurrent supported sessions vary with different data environments and also depend on the other type of work loads that the storage appliance has to support in addition to the backup/restore.

For more information on sizing and best practices on various NetApp data protection applications, refer to the appendix.

5.2 Dealing with Different Data Types

Data type, and how compressible it is, has a significant effect on backup and recovery rates. For this reason, whenever possible separate data that is highly compressed, such as video streams or images, from data that is not compressed, such as text files or source code. Segregate data like this into separate volumes or qtrees. Then, for volumes or qtrees with mostly compressed data, use a tape device alias that does not compress in hardware. For volumes or qtrees with mostly uncompressed or compressible data, use a tape device alias that does compress in hardware. This will improve overall performance in backing up and restoring from tape.

5.3 CPU Load Concerns

CPU load has a slight impact on backup and recovery performance once the load reaches a certain point. In addition, backup and recovery processes will themselves slightly increase the CPU load. Running multiple simultaneous backup or restore processes will compound this effect. As the number of simultaneous backup or restore processes is increased, it will eventually begin to affect user performance on a loaded system.

For optimum backup and recovery performance, avoid backup and recovery operations during times of operation when CPU load is over 80%. Also, run only as many concurrent backup or recovery operations as possible without driving the CPU load over 75%.

NetApp storage appliances' CPU performance can be monitored using DFM's performance advisor or alternatively using the Microsoft® MMC performance monitor (perf mon) plug-in.

5.4 Tape Drive Performance

Tape drive performance specifications play a very important role in backup-and-restore performance. Higher-performance tape drives such as Sony DTF-2 drives or IBM Ultrium LTO drives will result in shorter backup and restore sessions over slower drives such as Quantum DLT 7000 drives. This subsection provides an overview of supported tape drive uncompressed and compressed performance, as stated by the tape drive manufacturer.

MANUFACTURER AND MODEL	SUSTAINED NONCOMPRESSED TRANSFER RATE (MB/S)	SUSTAINED COMPRESSED TRANSFER RATE (MB/S)
Exabyte 8900 / Mammoth-1	3	6
Exabyte Mammoth 2	12	30
HP LTO 230	15	30
IBM 3950B	9	27

IBM 3590E	14	34
IBM Ultrium LTO / 3580	15	30
Quantum DLT 4000	1.5	3
Quantum DLT 7000	5	10
Quantum DLT 8000	6	12
Quantum SDLT 220	11	22
Quantum SDLT 320	16	32
Seagate LTO	15	32
Sony AIT-1	3	6
Sony AIT-2	6	12
Sony AIT-3	12	24
Sony DTF	12	20
Sony DTF-2	24	40
StorageTek 9840A	10	20
StorageTek 9840B	19	38
StorageTek 9940A	10	20
StorageTek 9940B	30	70
LTO2	30	60
LTO3	80	160

Table 4) Supported tape drive transfer speeds.

A detailed matrix of supported tape drive types can be found on the NOW™ site at <http://now.netapp.com/>.

5.5 Tape Devices

Below are some recommendations pertaining to tape devices:

- Attach tape devices to servers where large amounts of data reside
- For best performance, run concurrent backups to the maximum number of local attached tape drives supported
- Use tape drives in a tape stacker or library. Multitape backups on standalone tape drives require manual intervention
- Use a full-featured NDMP-compliant backup and recovery solution with media management to manage tape libraries

5.6 Tape Drive Recommendations

- For the best tape backup rates on large systems, use the maximum number of tape drives and run local backups concurrently

- Some tape drive models can be purchased with either fast/wide/single-ended (F/W/SE), fast/wide/low-voltage differential (LVD), or fast/wide/high-voltage differential (F/W/D or HVD) interfaces. We recommend that customers use the LVD or HVD interfaces because the NetApp storage system can be placed up to 12 meters (LVD) or 25 meters (HVD) away from an external tape drive, versus 6 meters for F/W/SE. The 6-meter limit is not very practical with external SCSI devices. Some tape drives are available with native Fibre Channel (FC) interfaces, which when used with FC switches can be up to 2 kilometers away

There are faster tape drives available these days. LTO Ultrium provides higher quality, better performance, and a better TCO than proprietary formats. For example, the LTO-3 has a compressed capacity of 800GB and a native capacity of 400GB with an Ultra160 SCSI interface. However, there are times when a virtual tape library is the way to go.

Although tape capacities and tape drive performance are improving, they are not improving at a rate fast enough to keep pace with advances in disk technologies and the amount of data being stored on typical networks. When this is coupled with a shortage of skilled IT resources and the ever-increasing demand for high data availability, the burden that backups place on an organization is immense.

In order to address the above mentioned concerns, NearStore Virtual Tape Library (VTL) can be used to streamline the tape backup process, improve performance, and shrink backup windows. NetApp VTL holds a significant role in the data protection strategy recommended by NetApp.

6. Summary of Recommendations

This section summarizes the Network Appliance recommendations for data protection.

6.1 Choosing a Backup Software Technology

Organizations protect data so that it can be restored when needed. Data recovery falls into three categories:

- Recovery of accidentally deleted files
- Long-term single or multiple file recovery from archived data
- Recovery of a file system after a disaster

Network Appliance provides a range of solutions that address these needs.

Snapshot Technology for Online Recovery from User Errors

The Data ONTAP Snapshot technology allows fast and easy recovery of accidentally deleted files. By scheduling Snapshot copies throughout the day, a system administrator can guarantee that recent files are available for recovery. Users can easily copy data from a Snapshot directory to their own directories. Because the Snapshot technology is designed into the system, the overhead is minimal.

FlexVol

FlexVol technology delivers true storage virtualization solutions that can lower overhead and capital expenses, reduce disruption and risk, and provide the flexibility to adapt quickly and easily to the dynamic needs of the enterprise. FlexVol technology pools storage resources automatically and enables you to create multiple flexible volumes on a large pool of disks.

FlexClone

FlexClone technology enables true cloning—instant replication of data volumes and data sets without requiring additional storage space at the time of creation. Each cloned volume is a transparent, virtual copy that can be used for essential enterprise operations, such as testing and bug fixing, platform and upgrade checks, multiple simulations against large data sets, remote office testing and staging, and market-specific product variations.

SnapVault for Disk-to-Disk Backup and Reduced Impact on Operations

SnapVault software provides a fast and efficient means for performing disk-to-disk backup and recovery for Network Appliance storage systems and open systems platforms. Since only block-level incremental changes are stored on the destination, but are usable as full backups, space utilization is minimal and restores are very fast and easy to perform. NDMP-based tape backup can then be used to protect this data on the destination, offloading all overhead from tape-based data protection from primary storage systems.

SnapMirror for Disaster Recovery

For disaster recovery, Network Appliance recommends the use of SnapMirror to mirror data to a remote location. In case of serious file system damage, the system administrator can turn the mirrored file system into the active file system. Even if a disaster occurs at one site, the data remains safe and online in the remote location.

MetroCluster

MetroCluster extends failover capability from within a data center to a site located many miles away. It also replicates data from the primary site to the remote site to ensure that data there is completely current. The combination of failover and data replication ensures recovery from disaster—with no loss of data—in minutes rather than hours or days. The built-in simplicity of MetroCluster allows for near continuous storage service operations during disasters with little to no administrator intervention required.

Tape Backups for Archiving Data

For archive purposes, Network Appliance recommends customers use one of the NDMP-compliant third-party solutions to back up data to tape. Tape is currently the most reliable removable media. Furthermore, the third-party data protection solutions can create indexes of data so that the data can be recovered in the future. NDMP invokes the dump/restore primitives, which have been optimized in the Data ONTAP kernel for both performance and low system impact.

NearStore VTL

In order to address the issues related to backup windows and performance, NearStore Virtual Tape Library (VTL) emerges to be a faster, more cost effective, more reliable, and easier-to-manage replacement for tape libraries. The NearStore VTL device enables a disk storage system to look exactly like one or more tape libraries to the backup application. It is the easiest way to incorporate disks into a backup environment. This enables customers to rapidly deploy disks for their backup and restore needs and begin enjoying the benefits of disks—faster recovery times and simplified management—immediately. At the same time, this VTL solution integrates with real physical tape libraries for the seamless creation of removable media sets.

SnapLock

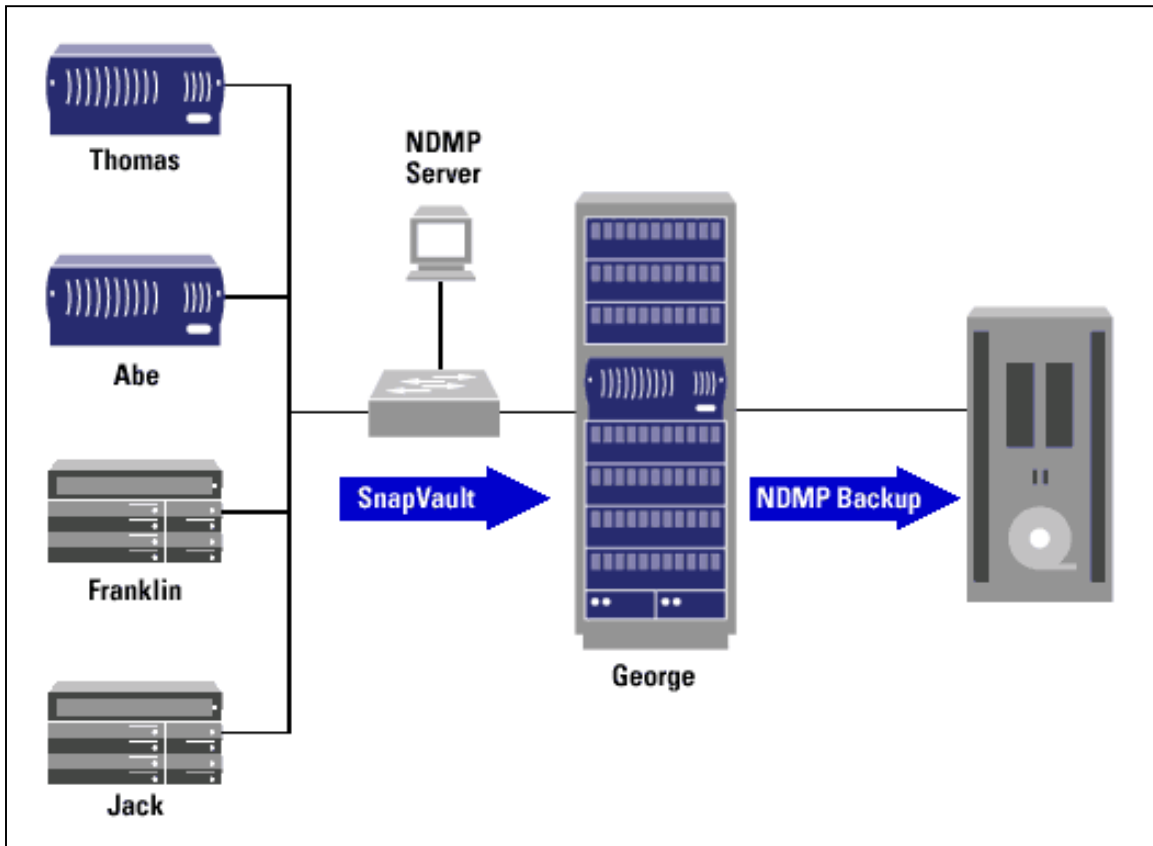
NetApp SnapLock enables compliance with regulatory and best-practices records-retention requirements by allowing the creation of nonrewritable, nonerasable WORM volumes on NetApp NearStore and FAS storage systems, thereby preventing critical files from being altered or deleted until a specified retention date.

LockVault

LockVault allows customers to make permanent WORM backups of their unstructured data, by simply copying and storing only the unique blocks that have been written since the most recent incremental backup. Every block variation is protected, so no data falls through the cracks. LockVault also automatically provides online compliant backups of unstructured data at multiple points in time.

6.2 Configuring NearStore as a Backup Storage System

To gain full advantage of Network Appliance data protection technologies and methodologies, you can configure NearStore as a backup storage system, as shown in Figure 4. In this configuration, the Network



Appliance storage systems, Thomas and Abe, and the open systems servers, Franklin and Jack, replicate their data to four discrete volumes on the NearStore device, George. An NDMP-compliant backup solution can then run local simultaneous tape backups from NearStore to a high-capacity, multidrive tape library. The steps are as follows:

1. Run a SnapVault baseline transfer for each of the four systems
2. Run incremental updates to the volumes according to the schedule that fits your needs
3. Back up the four volumes on George in parallel to multiple tape drives in the tape library

Steps 2 and 3 can overlap. For example, you can run incremental SnapVault transfers on Thomas and Abe while backing up the volumes Franklin and Jack to tape, creating a pipeline schedule to fit your environment.

Figure 4) Network Appliance NearStore as backup storage system.

This configuration offers the following benefits:

- Allows for tape backups with no backup window
- Reduces load on application systems because they are not running local tape backup sessions
- Meets archiving needs
- Allows load balancing with read-only access of data on NearStore
- Fits into existing NDMP-based third-party infrastructures

6.3 Organizing File System Data into Volumes and Quota Trees

NetApp storage systems support multiple volumes on the same storage system and multiple qtrees within each volume. Backup performance can be optimized by setting up volumes and qtrees correctly. Network Appliance recommends the use of multiple volumes to ensure that the volume size meets your recovery time estimates. Within volumes, configure qtrees. The reasons follow:

- A full dump works on an entire volume. A full dump of a very large volume will take longer than a full dump of a smaller volume. You can minimize backup windows by dividing your system into smaller volumes and doing full dumps of different volumes on different nights.
- Large volumes take longer to restore. Match volume size to an acceptable time for a full disaster recovery from tape. For example, if your restore window for any volume is eight hours and you estimate restore rates to be 40GB per hour, you can use the following formula to calculate what your largest volume should be:

$$40\text{GB/hr} * 8 \text{ hours} = 320\text{GB volume}$$

If a 12-hour restore window is acceptable, the volume size can go up to 480GB.

RAID-DP, double-parity RAID, is a unique NetApp technology that provides protection against data loss equivalent to RAID1 (mirroring) without the impact on usable capacity. RAID-DP groups can be configured large enough to match RAID4 parity overhead (i.e., cost impact), and performance is fundamentally the same as RAID4. RAID-DP is available with no cost or special hardware requirements. The only requirement to start using RAID-DP is to upgrade to at least Data ONTAP version 6.5.

Traditional single-parity RAID technology offers protection from a single failed disk drive. The caveat is that no other disk fail or that uncorrectable bit errors not occur during a read operation while reconstruction of the failed disk is still in progress. If either secondary event occurs during reconstruction, then some or all data contained in the RAID array or volume could be lost. With modern larger disk media, the likelihood of an uncorrectable bit error is fairly high, since disk capacities have increased but bit error rates have stayed the same. Hence the ability of traditional single-parity RAID to protect data is being stretched past its limits. The next level in the evolution of RAID data protection is double-parity RAID, or RAID-DP, available on the entire Network Appliance data storage product line.

6.4 Classifying Data by Value and Change Rate

We recommend that customers classify data by its value and by how dynamic it is. Isolating different types of data into different volumes or qtrees allows flexibility in backup policies, for both tape backups and Snapshot copies. Important and dynamic data in one volume may warrant frequent full dumps to tape. Infrequently changing archival data can be placed in a separate volume, which gets occasional incremental backups and full dumps only at widely spaced intervals.

Multiple volumes or qtrees also allow you to configure different Snapshot schedules for different kinds of data. For example, you may have home directories and a source code repository on the same NetApp storage system. Enabling frequent Snapshot copies for home directories so that users can always restore their own files saves time and system administrator resources. One would disable Snapshot copies for the source code repository because it changes so frequently; therefore, Snapshot copies are not practical.

7. Summary

Risks of data loss include deleted files, system crashes, application crashes, viruses, and natural disasters. Because these risks abound, a carefully thought out data protection plan is imperative. Putting a data protection plan in place involves:

- Identifying business-critical data
- Defining requirements such as:
 - Protecting from accidentally deleted files
 - Archiving data for future use
 - Reducing backup and restore windows
 - Recovering from a disaster
- Analyzing your data as to how dynamic it is, total data set size, number and size of files, directory structure, and data types
- Choosing data protection software and hardware solutions that address your requirements
- Configuring these technologies so performance is adequate for your needs

Network Appliance provides a unique set of solutions. Built-in Snapshot technology, enabling online Snapshot copies throughout the day, provides online file recovery plus a potential replacement for time-consuming incremental tape backups. SnapRestore software allows near-instantaneous recovery of individual files or entire volumes, minimizing downtime to recover from virus infections or database corruption. NDMP gives NetApp users a variety of choices among high-performance tape backup and restore vendors. The SnapMirror technology enables organizations to replicate data volumes at high speeds over a network, providing an up-to-date mirrored volume to ensure uninterrupted operation in case of disaster. And SnapVault software enables efficient disk-to-disk backup and recovery of both NetApp storage systems and open systems storage. The best data protection strategies often combine these solutions, each of which solves a different piece of the backup problem.

8. Revision History

Status	Date	Author
Update	March 2007	Pradeep Seshadri
Update	July 2006	Pradeep Seshadri
New	October 2005	Gilda Farvid, Nicholas Wilhelm-Olsen, Jay Desai, Grant Melvin, Mike Federwisch

Appendix

Some useful links to various data protection best practice guides:

- SnapMirror Best Practices Guide
www.netapp.com/library/tr/3446.pdf
- OSSV Best Practices Guide
www.netapp.com/library/tr/3466.pdf
- SnapVault Deployment and Configuration Guide
www.netapp.com/library/tr/3240.pdf
- Using SnapLock with Data ONTAP 7G
www.netapp.com/library/tr/3342.pdf
- Storage Best Practices and Resiliency Guide
www.netapp.com/library/tr/3437.pdf
- Synchronous SnapMirror Design and Implementation Guide
www.netapp.com/library/tr/3326.pdf

NDMP certification matrixes of third-party backup software:

- Atempo Time Navigator
www.atempo.com/fileadmin/user_upload/files/en/CompatibilityGuide_TN40.pdf
- Bakbone NetVault
www.bakbone.com/docs/Supported_NAS_Appliances_and_NDMP_Compatibility_Guide.pdf
- Computer Associates BrightStor ARCserve Backup r11.5
supportconnectw.ca.com/public/storage/infodocs/basb115win/basb115win_cdl.asp#nas_ndmp
- HP OV Data Protector 5.5
www.managementsoftware.hp.com/products/storage_data_protector/device_matrices/NAS_Support_Matrix_DP55.pdf
- Syncsort Backup Express 2.35
www.syncsort.com/products/bex/partners/technologypartners/partners/netapp.htm
- Symantec NetBackup
http://ftp.support.veritas.com/pub/support/products/NetBackup_DataCenter/251713.pdf

NearStore VTL compatibility guide:

http://mktg-web.corp.netapp.com/products/hardware/vtl/nearstore_vtl_compatibility.pdf



www.netapp.com

© 2006 Network Appliance, Inc. All rights reserved. Specifications subject to change without notice. NetApp, the Network Appliance logo, DataFabric, FAServer, FilerView, NetCache, NearStore, SecureShare, SnapManager, SnapMirror, SnapRestore, SpinCluster, SpinFS, SpinHA, SpinMove, SpinServer, and WAFL are registered trademarks and Network Appliance, ApplianceWatch, BareMetal, Camera-to-Viewer, ContentDirector, ContentFabric, Data ONTAP, EdgeFiler, HyperSAN, InfoFabric, MultiStore, NetApp Availability Assurance, NetApp ProTech Expert, NOW, NOW NetApp on the Web, RoboCache, RoboFiler, SecureAdmin, Serving Data by Design, Smart SAN, SnapCache, SnapCopy, SnapDirector, SnapDrive, SnapFilter, SnapMigrator, Snapshot, SnapSuite, SnapVault, SohoCache, SohoFiler, SpinMirror, SpinShot, SpinStor, The evolution of storage, Vfiler, VFM, Virtual File Manager, and Web Filer are trademarks of Network Appliance, Inc. in the U.S. and other countries. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.