TechnicalReport

Data Protection Handbook

Data Protection Technical Marketing, NetApp

TR-3784

Version 1.1

ABSTRACT

The purpose of this document is threefold. Volume 1 discusses the challenges that organizations face today in protecting their data and some of the strategies that they use to do so. Volume 2 goes a step further, discussing how the various NetApp® products, either alone or in combination, can meet those increasingly demanding data protection needs. Volume 3 provides example scenarios of how these solutions are implemented.

TABLE OF CONTENTS

1 STRATEGIES FOR DATA PROTECTION

1.1 INTRODUCTION

Electronic data is critical to the success and survival of modern businesses. Companies understand that the amount of data they rely upon is increasing no matter what the condition of the economy and the business climate. Faced with the facts that data is becoming more and more important and also that there is more and more of it, organizations are continually challenged to protecting that data.

Why keep so much data and add to the challenge every year?

First, organizations may not have a choice. Current regulatory requirements often force them to keep data for extended periods. Another reason is that accumulating more data enables more effective analysis and better business decisions. This often becomes a competitive advantage. So what is the best way for an organization to protect its data? The answer is the magic phrase “It depends!” The purpose of this handbook is to answer the inevitable follow-on question, “Depends on what”?

The NetApp Data Protection handbook is broken in to three sections:

Volume 1 discussesdata protection fundamentals. It answersquestions like:

- What exactly is data protection?
- What do all the sometimes confusingdata protection terms mean?
- What exactly are we protecting data from?
- Does all data need protection and at the same level?
- Whatstrategies can we employ not only to protect data but to recover it as well?

Volume 2 outlines how NetApp products can be used to meet protection requirements and to implement the chosen strategies.

Volume 3 offers specific implementation examples on how to safeguard a corporation’s data to enhance its chances for survival and success.

1.2 WHAT IS DATA PROTECTION?

Data protection is the act of safeguarding important data from damage, alteration, or loss. Although that definition sounds simple and obvious, it encompasses a host of technologies, business processes, and best practices. Different techniques must be used for different aspects of data protection. For example, securing the storage infrastructure so that data is not altered or maliciously destroyed may not use the same techniques as those used to protect against inadvertent data loss or permanent corruption.

A variety of considerations, from the importance of the data to budget and person power, determine which practices, processes, or technologies are used for data protection. For instance, in most cases it is not reasonable to assume that a small business can deploy expensive, high-end solutions to protect important data. On the other hand, backing up data to tape or disk is certainly something that any enterprise can do. A large enterprise typically hasboth the resources and the motivation to use more advanced technology.

No matter what the size or makeup of the company,the goal of data protection is the same—to minimize business losses due to the lack of verifiable data integrity and availability.

1.3 NAVIGATING THE TERMINOLOGY

If youwere to ask 15people the following questions, you would probably get 15 different answers:

- What is the difference between business continuity and disaster recovery?
- What is the difference between high availability and disaster recovery?
- Do I need backup and recovery or disaster recovery?

To cut through the confusion, it's better to think less about specific terms and to focus on the data protection strategy necessary for a given situation or type of application and data. In case you're interested, though, a glossary is provided at the end of Volume 1.

BUSINESS CONTINUITY VERSUS DISASTER RECOVERY

Business continuity encompasses everything a business needs to continue its operations before, during, and after a disaster. This includes facilities, staff, communications, information, and IT resources, among other things. Business continuity is not limited to IT resources; it includes all resources necessary to continue operation.

Disaster recovery (DR) is actually a subset of business continuity; it applies to everything that is necessary for a business to recover from a disaster. DR strategies encompass all of the tasks necessary to get core business functions backup and running so the business can operate, even if it is in degraded mode, until the rest of the business operations are online. Whereas a business continuity plan includes getting all operations back up and running (for example, sales, engineering, support, and operational staff like HR), a DR strategy focuses on getting the lights back on and key production operations up.The time to complete both business continuity and disaster recovery operations can vary from minutes to hours to days, depending on the criticality of specific business requirements.

It's worth exploring at this point what constitutes a disaster. The most common definition is"a critical event that affectsthe ability of a business to continue operations at a specific location or locations."This includes earthquakes, hurricanes, and other natural and human-causeddisasters that eliminate power, connectivity, and overall operational capabilities. This is why disaster recovery implies an alternate location or DR site. The reality is that in a small company with their entire ERP database on a shelf of disk drives, ifsomeone kicks out a power plug or their database becomes corrupted, this it's no less a disaster than if an earthquake occurred. Either event affectsthe business and requires recovery in terms of business operations and data.

HIGH AVAILABILITY VERSUS DISASTER RECOVERY

High availabilityencompasses technologies and business practices that significantly limit or eliminate the impact of outageson business operations. High availability can occur at a component level, within a system, within a data center (for example, via clustering), or across sites. A good example is redundant components that can take over for those that fail without disrupting operation. The net is that high-availability strategies are designed for almost instantaneous recovery and are often designed on an application or business process level, which may not account for full site outages. Disaster recovery tasks, on the other hand, focus on much more macro events and can have a longer outage period because of the effort required to restore many applications and business processes. So when it comes to high availability, is it one orthe other? For operations that are most critical to the business,it’s both. Business-critical applications require minimal disruptions with quick recovery when an outage does occur.

BACKUP AND RECOVERY VERSUS DISASTER RECOVERY

Most people think of backup and recovery in localized terms, including the restoration of accidently deleted files, corrupt databases, and local system failures. Backup and recovery is built around the requirement to keep multiple point-in-time copies of data sets available so that an organization can roll back to the correct “view” of the data, before data loss or corruption occurred. Disaster recovery tends to involve outages that affect an entire site and, although it often relies on just the most current revision of a data set, it may require rollback to a previous version of data to avoid using corrupted or incomplete data sets. This sometimes leads to the question, “Can a backup and recovery solution also be a disaster recovery solution?”The answer is yes. For most organizations, backups—whether they are performed via traditional backup applications or snapshots—are used to create the offsite data copy used for restoration in the event of a disaster. That said, backup and recovery can be independent of disaster recovery operations, depending on recovery point objectives (discussed later in this section), because backup usually happens once per day, whereas critical disaster recovery operations may occur continuously.

RECOVERY TIME OBJECTIVE (RTO)

A recovery time objective is the period of time within which systems, applications, or functions must be recovered after an outage. This can be measured in minutes, hours, or days, depending on the criticality of the resource. Do all resources have the same RTO? The answer is no. If the primary data center goes down, access to business-critical databases tends to be far more important than users’ home directories. Table 1 shows the average recovery times broken down by industry.

Table 1-1) Recovery time objectives by industry.

	Less than 4 hours	4 - 12 Hours	13-24 hours	25-72 hours	More than 72 hours
Utilities			68.70%	33.30%	
Transportation	45.00%	25.00%	10.00%		20.00%
Services	27.30%	21.80%	25.50%	21.80%	3.60%
Retail/Wholesale	17.20%	6.90%	20.70%	24.10%	31.00%
Local/State/Regional Government	73.30%	13.30%	6.70%	6.70%	
Federal Government	33.30%	8.30%	8.30%	33.30%	16.70%
Healthcare Provider	37.10%	17.10%	11.40%	22.90%	11.40%
Financial Services	32.30%	31.30%	14.60%	18.80%	3.10%
Energy		66.70%	33.30%		
Education	12.20%	12.20%	18.40%	36.70%	20.40%
Manufacturing	13.30%	18.30%	16.70%	16.70%	35.00%
Communications	32.10%	32.10%	10.70%	21.40%	3.60%
Average	32.31%	23.00%	16.03%	22.49%	16.09%

Source: Glasshouse

RECOVERY POINT OBJECTIVE (RPO)

A recovery point objective represents the allowed age of the data when restoration occurs. For instance, most backup strategies have a recovery point objective of 24 hours (the data cannot be older than 24 hours). RPOs are assigned on a data criticality level and come down to how much recently created or modified data an organization can tolerate losing (Figure 1).Shorter RPOs can be achieved by using continuous or near-continuous data protection technologies. It is also interesting to note that certain data sets can have multiple RPO requirements (for example, an order management database is replicated so that the RPO is near-zero data loss as well as backed up so that RPOs of 24 hours, 48 hours, and so on are available in case of a corruption event and the need to roll back to a previous state. Determining the proper RPO levelrequires an assessment along with RTO requirements.

.

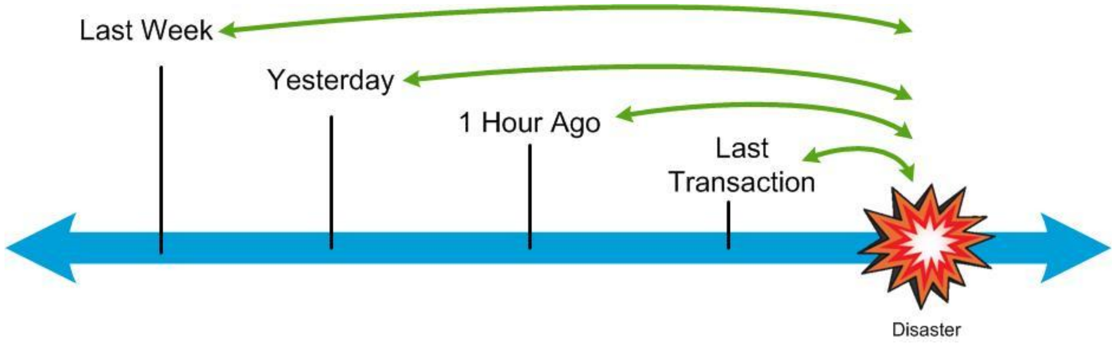
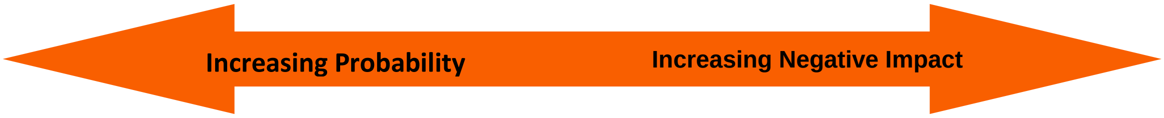


Figure 1-1) Recovery point objectives.

1.4 THREATS TO YOUR DATA

From what kind of threats does an organization need to protect its data? As shown in Figure 2, there are many causes of failures and disasters. About 80% of the failure and outages are due to operational and application failures. Many organizations focus all of their data protection in these areas, figuring that if they can protect against 80%, they're in good shape. The problem with this emphasis is that the other 20%,although less likely to occur,have a much greater impact on business operations or even on the survival of the business. The key to a thorough data protection strategy is to employ technology and products that can scale across the spectrumof protection needs.



Failures and Disasters	Operational Failures	Application Failures	Component Failures	Data Center Failures	Regional Disasters
	40%	**40%**			
Causes	People Issues	Bugs	Controller	Terrorism	Power Grid
	Process issues	Performance Issues	HBA Port Failures	Cyber Threats	Floods
	Infrastructure Changes	Change Management	Disk Failure	Fire	Hurricanes
	Configuration and Problem Management	Upgrades	Shelf Failure	Plumbing	Earthquakes
			FC Loop Failure	Architectural	
				Planned Downtime	

Figure 1-2)– Causes of failures and disasters.

THE HUMAN THREAT

The word “threat” sounds ominous and seems to imply malicious intent, but the fact is that many risks to data integrity and access are caused either directly or indirectly by humans, not always maliciously. It could be due to lack of training or experience or the simple common characteristic all humans share—they make mistakes. The issues are whether data access or integrity is affected (or both) and how critical the recovery is. A more proactive way of looking at human threats is how they can be prevented. Proper training is one key. Another is the complexity of the administrative tasks and how many different tools are required to manage the environment. The simpler things are to manage and the fewer tools necessary, the less the chance of a human error occurring.

ENVIRONMENTAL FAILURES

Environmental failures can range from those within the data center, such as air conditioning and power failures, to regional natural disasters such as earthquakes, hurricanes, terrorist acts, and tornados. Local problems can be addressed by redundant power grids, air conditioning systems, and even alternate data centers across campus. Several recent surveys state that the most common cause of data center failure is power issues. Although there is only so much that can be done in terms of protection, it’s still important to plan for these types of failures in terms of contingencies for critical business applications.

HARDWARE FAILURES

Just as humans make mistakes, the fact is that hardware fails. Even though the quality of hardware and the time between failures has improved over the years, failures still happen.

Most systems that house critical data have redundant cooling, power supplies, network cards, and controllers in order to eliminate or minimize downtime due to component failure.

In addition to redundant critical components, other technologies have been developed and implemented, including RAID configurations and write protection. A redundant array of independent disks (RAID) provides the ability to recover from the loss of one or more disk drives, in some cases even providing better performance. There are many different configurations, the details of which are beyond the scope of this paper, but the more disk failures that can be recovered from, the better. Disks are among the most mechanical devices that house data and are therefore the most prone to wear and failure.

Another technology that is increasingly important is write protection, which involves the ability to keep data that has been written from being erased or altered for either a set period of time or forever. Legal, medical, and insurance information are prime candidates for write protection. Also, an increasing number of government regulations require the write protection of data for a certain number of years. Therefore it’s important to consider the use of write protection in developing a data protection strategy.

SOFTWARE FAILURES

Software is also prone to failures. Because software creation is a human endeavor, it is subject to flaws. Some flaws affect access to data, while others can affect the integrity of the data itself. As a result, software upgrades are commonplace, usually requiring downtime to perform. Sometimes the upgrades actually introduce new problems.

The keys to mitigating risk are to have backup copies of the data for restore and also to be able to nondisruptively test software upgrades before moving them into production. The ability to perform the upgrade nondisruptively is an added bonus.

1.5 DATA PROTECTION AND RECOVERY METHODS

Now that we have discussed the numerous threats to access and integrity of business-critical data, it’s time to talk about making the correct choices for a data protection strategy. The only way to choose the correct protection methods is to analyze your organizations requirements.

ANALYZING YOUR REQUIREMENTS

In analyzing your organization’s data protection requirements, it’s important to consider obvious tradeoffs. Protecting an organization’s data involves cost. How much cost? It depends on the level of protection required, which in turn depends on the value of that data to the business. As you can see, there are tradeoffs. The key is to properly analyze the data so that the correct protection can be achieved in the most cost-effective manner.

Business criticality

The first step in determining criticality is to perform an inventory and categorization of business applications and data,summarizing this information in a table similar to Table 1-2.

Table 1-2) Business criticality evaluation.

Attribute	Mission Critical 15%	Vital 20%	Sensitive 25%	Noncritical 40%
Recovery Time Objective	Immediate	Seconds	Minutes	Hours, Days
Availability Index	99.999+	99.99	99.9	<99.9
Retention Period	Hours	Days	Years	Infinite

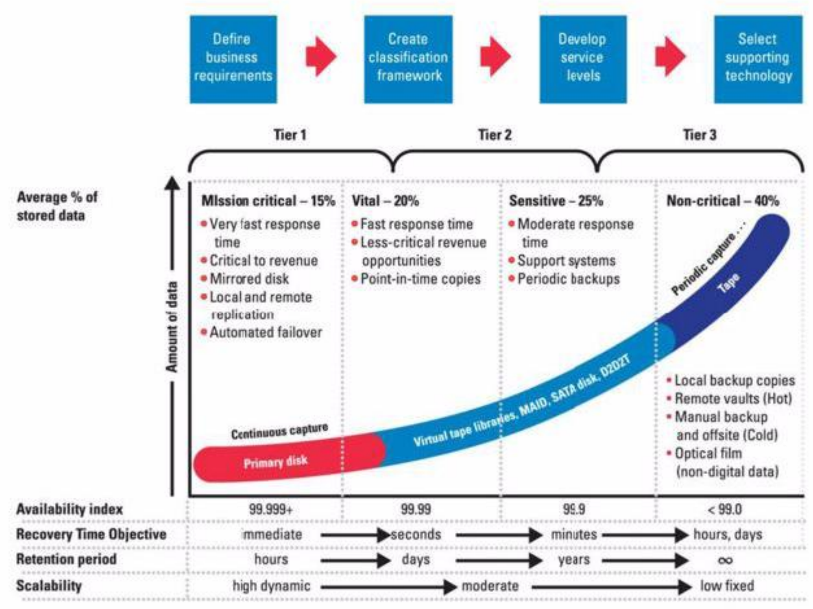
Source: Horizon Information Strategies

Regulatory requirements

Some of the choices for data protection methods and strategies are dictated by government regulatory agency mandates. The important point is to understand exactly what those requirements are and to what kind of data they apply. Most of the regulatory requirements center around data retention and the ability to protect data records from modification. The ability to keep data from being modified either for a certain period of time or indefinitely is extremely important.

Data classification

An essential step in determining a company's data protection strategy is to identify the criticality of the data in the enterprise, and how quickly the data must be recovered in case of disaster. You must also decide how current your recovered files must 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. Figure 1-3 shows the data classification model hierarchy of critical data. The value of the information dictates the technology you use to protect it.



Source: Horison Information Strategies

Figure 1-3) Dataclassification model.

Change rate

One critical piece of information that organizations need as they put together a disaster recovery plan is how much data they have in their environment and how quickly it is changing. This information is so important because without it, organizations often have no way to effectively size how much or what type of capacity they need to protect and recover their production and archive data.

The other sizing consideration affected by change rate is the network bandwidth required whenbacking up or replicating those changes over a network. The problem in using this metric is that it doesn't necessarily tell you when the changes are occurring. For example if you have 50TB of data and 5TB of it changes each day (10% change rate), you still don't know if those changes occurred evenly spread over 24 hours or within a short period. If you are using periodic snapshot copies, then it is useful to measure the difference between each of the copies over that 24-hour period.

Also, remember not to use change rate only in storage capacity planning because it alone does not tell you how much data is new versus changed data.

Frequency of backup or replication

How often does the data need to be replicated or backed up? That depends on factors such as recovery point, data classification, and backup or replication methodology. It also depends on the backup window. This is where it is determined when backups need to be performed and how long they should take. However, there may not be an ideal time period. Some applications and services must be available 24 hours a day, 7 days a week. In these cases, you must use other methods of obtaining a consistent backup. In other situations, the time required to perform a backup may exceed the time of server inactivity, so consider the following compromises:

- Decrease the frequency of backups.
- Accept a lower level of data availability.
- Accept a decrease in application performance.

Backup frequencies depend primarily on how much the data changes and its importance. Businesses typically do weekly full backups on the weekends and nightly incremental backups Monday through Friday. The hardcore approach is to do a full weekly backup, duplicate it, store the copies in separate locations (different cities or states is even better), and then do nightly incremental backups. That way, if something catastrophic happens (such as the destruction of the building with one set of backups), the business is never more than one week out of date.

Data retention

Organizations these days need to keep their data for longer periods of time. Several factors can affect this decision. As mentioned earlier, regulatory requirements may dictate that certain types of data are retained for a minimum period of time.

Another factor that can affect retention period is resource limitations. Organizations may not be able to afford large-scale tape libraries or massive quantities of disk storage. This is where the efficiency of the backup or replication can really assist. If the data can be deduplicated (removal of redundant data), then more data can be retained for a longer period in a given quantity of storage.

Also, from an efficiency standpoint it's important to look at each type of data in terms of retention. Not all data needs to be retained forever (although disk and tape storage vendors may disagree).

Recovery frequency

Many organizations look at data protection strictly from a backup perspective and don't consider how often or how quickly data needs to be recovered.As with most of these items,these needsmay vary with the classification of data. That's why an organization may need to consider more than one backup and recovery method. With user home directories, the need to recover accidentally deleted files may happen more than a few times each week. Itwould make life much easier forboth the user and the administrator if userscould recover the data themselves.

SUMMARY

There are many factors to consider when looking at data protection methods and developing a strategy. The decisions to be made are more complex becausethese factors are often interrelated. Sometimes tradeoffs have to be made. That is why it's so important is to complete a protection requirements summary along the lines of that shown in Table 1-3.

Table 1-3) Data protection requirements summary.

Application	Oracle	Exchange	Home Dirs	Other
Data Type				
Data Classification				
RPO				
RTO				
Change Rate				
Regulatory Requirement				
Backup or Replication Frequency				
Retention				
Recovery Frequency				

In documenting these requirements, it is also vital to consult the actual owners and users of the data. There have been many occasions whendata protection requirements developed without this consultation resulted in insufficient protection, which typically isn’t discovered until a problem occurs.

GENERAL PROTECTION METHODS

Hardware

"RAID" is an umbrella term for computer data storage schemes that divide and replicate data among multiple hard disk drives. The various RAID designs all involve two key goals: increased data reliabilityand increased input/output performance. When multiple physical disks are set up to use RAID technology, they are said to be in a RAID array. This array distributes data across multiple disks, but the array is seen by the computer user and the operating system as one single disk. RAID can be set up to serve several different purposes.

Redundancy is achieved either by writing the same data to multiple drives (known as mirroring) or by writing extra data (known as parity data) across the array, calculated so that the failure of one (or possibly more, depending on the type of RAID) disks in the array does not result in loss of data. A failed disk can be replaced by a new one and the lost data reconstructed from the remaining data and the parity data. Organizing disks into a redundant array decreases the usable storage capacity. For instance, a two-disk RAID 1 array loses half of the total capacity that would otherwise have been available using both disks independently, and a RAID 5 array with several disks loses the capacity of one disk. Other types of RAID arrays are arranged so that they are faster to write to and read from than a single disk.

Various combinations of these approaches give different trade-offs of protection against data loss, capacity, and speed. RAID levels 0, 1, and 5 are the most common, and they cover most requirements.

- RAID 0 (striped disks) distributes data across several disks in a way that gives improved speed and no lost capacity, but if any one disk fails, all data on all disks is lost.
- RAID 1 (mirrored settings/disks) duplicates data across every disk in the array, providing full redundancy. Two or more disks each store exactly the same data, at the same time, and at all times. As long as one disk survives, no data is lost. Total capacity of the array equals the capacity of the smallest disk in the array. At any given instant, the contents of each disk in the array are identical to that of every other disk in the array.
- RAID 5 (striped disks with parity) combines three or more disks in a way that protects data against loss of any one disk; the storage capacity of the array is reduced by one disk.
- RAID 6 (striped disks with dual parity; less common) can recover from the loss of two disks.
- RAID 10 (or 1+0) uses both striping and mirroring. "01" or "0+1" is sometimes distinguished from "10" or "1+0": a striped set of mirrored subsets and a mirrored set of striped subsets are both valid, but distinct, configurations.

RAID can involve significant computation when reading and writing information. With traditional "real" RAID hardware, a separate controller does this computation. In other cases, the operating system (or simpler and less expensive controllers) requires the host computer's processor to do the computing, which reduces the computer's performance on processor-intensive tasks. Simpler RAID controllers may provide only levels 0 and 1, which require less processing.

RAID systems with redundancy continue working without interruption when one (or possibly more, depending on the type of RAID) disks of the array fail, although they are then vulnerable to further failures. When the bad disk is replaced by a new one, the array is rebuilt while the system continues to operate normally. Some systems have to be powered down when removing or adding a drive; others support hot swapping, allowing drives to be replaced without powering down. RAID with hot-swapping is often used in high-availability systems, where it is important for the system to remain running as much of the time as possible.

RAID is not a good alternative to backing up data. Data may become damaged or be destroyed without harm to the drive or drives on which it is stored. For example, part of the data may be overwritten by a system malfunction; a file may be damaged or deleted by user error or malice and not noticed for days or weeks; and, of course, the entire array is at risk of physical damage.

Write protection

WORM (write once, read many (WORM)storage refers to computer data storage systems, data storage devices, and data storage media that can be written to once and read from multiple times. This technology is particularlyimportant in meeting regulatory requirements for data protection.

WORM is an inherent property of some data storage media and devices, in which the storage medium is physically incapable of being rewritten. WORM storage media include CD-R and DVD-R.

AlthoughWORM can be an inconvenient property when it comes to reusing recorded media, it is a desirable property for data backups and archives, to prevent erasure (accidental or deliberate) and tampering. Various regulatory agencies require data such as health information and transaction records to be archived reliably and securely over a long period of time. Therefore WORM capability has been added to otherwise rewritable media such as magnetic tape data storage and hard disk drives. The media can be written to, but the written portion immediately becomes read-only.

Software solutions are more versatile, becausethey can allow additional data to be written to the media until it is full, while not allowing erasure or overwriting of data already written. But they also require physical control of the media. Furthermore, the software must be audited, and computer systems must be configured so that no other software can access the media.

Modern WORM systems overcome most problems by building a tested software solution into the cartridges, drives, controllers, or operating system. For removable media, the high-capacity drives are new enough that no available drive violates the rules. Other features make sure that any data somehow overwritten in spite of restrictions is detectable.

Storage systems thatcan restrict writing to one time only include Super DLT (Super DLT II is used for both standard and WORM operations, as specifiedby the user), Linear Tape-Open (LTO), AIT, and various hard disk drive-based systems. Some of these systems allow reuse of recorded media. A good disk-based system allows mixing WORM disks with other disks on the same system for increased flexibility. Efficient storage systems also allow WORM storage to be configured either permanently or for a specific period based on an expiration date.

Data can be destroyed by destroying the media, but the loss is visible. (The loss can be covered up by replacing the destroyed media with blank media.) Magnetic media can be erased or rendered unreadable by a strong magnetic field. Punched media can be invalidated by a few extra holes. Solid-state memory can be ruined by applying excess voltage. CD-R and DVD-R can be ruined by leaving them in sunlight or by applying microwave radiation.

BACKUP

Point-in-time copies

A point-in-time copy is typically a read-onlysnapshot of the file system that reflects selected data in a disk subsystem at a particular instance in time. Point-in-time copies are used for backups, availability, and disaster recovery solutions and can be created manually or on a scheduled basis. These copies offer protection from such events as accidental file deletions, data corruption, virus attacks, and so on.

-Recovery involves restoringthe affected data from the desired point-in time copy.

Physical tape

A tape drive, also known as a streamer, is a data storage device that reads and writes data stored on a magnetic tape. It is typically used for archival storage of data stored on hard drives. Tape media generally has a favorable unit cost and long archival stability. Physical tape backups, typically stored in a location separate from the production site, offer protection from most disasters at the production site. However, depending on frequency and type of backup, data since the last tape backup can be lost. Recovery involves restoring the affected data from one or more tapes, which can result in extremely lengthy recovery times.

Instead of allowing randomaccess to data as hard disk drives do, tape drives allow onlysequentialaccess of data. A hard disk drive can move its read/write heads to any random part of the disk platters in a very short amount of time, but a tape drive must spend a considerable amount of time winding tape between reels to read any one particular piece of data. As a result, tape drives have very slow average seek times. Despite the slow seek time, tape drives can stream data to tape very quickly. For example, modern LTO drives can reach continuous data transfer rates of up to 120MB/s and, using 2:1 compression, rates of up to 240MB/s.

Tape drives can range in capacity from a few megabytes to hundreds of gigabytes, uncompressed. In marketing materials, tape storage is usually discussed with the assumption of a 2:1 compression ratio, so a tape drive might be known as 80/160, meaning that the true storage capacity is 80 whilethe compressed storage capacity can be approximately 160 in many situations. IBM and Sony have also used higher compression ratios in their marketing materials. The real-world, observed compression ratio always depends on what type of data is being compressed. The true storage capacity is also known as the native capacity or the raw capacity.

In computer storage, a tape library, sometimes called a tape silo, tape robot, or tape jukebox, is a storage device thatcontains one or more tape drives, a number of slots to hold tape cartridges, a barcode reader to identify tape cartridges, and an automated method for loading tapes (a robot). These devices can store immense amounts of data, currently ranging from 20 terabytes up to more than 50 petabytes of data, or about 100,000 times the capacity of a typical hard drive and well in excess of capacities achievable with network-attached storage. Typical entry-level solutions cost around US$10,000, while high-end solutions can cost in excess of $70,000. For large datastorage, this is a cost-effective solution, with cost per gigabyte as low as $0.10, or at least 60% less than most hard drives and they also provide systematic access to very large quantities of data. The tradeoff for their larger capacity is their slower access time, which usually involves mechanical manipulation of tapes. Access to data in a library takes from several seconds to several minutes.

Because of their slow random access and huge capacity, tape libraries are primarily used for backups and as the final stage of digital archiving. A typical application of archiving would be an organization's extensive transaction record for legal or auditing purposes. Another example is hierarchical storage management, in which a tape library holds rarely used files from file systems.

Smaller tape libraries with only one drive and robot are known as autoloaders.

Disk-to-disk

The term disk-to-disk, or D2D, generally refers to disk-to-disk backup. With D2D, a computer hard disk is backed up to another hard disk rather than to a tape. D2D is often confused with virtual tape, but it differs in that it enables multiple backup and recovery operations to simultaneously access the disk directly by using a true file system, and it can accommodate multiple versions of data on the backup system. Typical advantages of disk-to-disk include:

- Higher speed and higher capacity, relative to tape, result in shorter backup and recovery windows.
- Nonlinear recovery of data enables a specific file to be restored faster and easier than with tape.
- Total cost of ownership is lower due to increased automation and lower hardware costs.

Disk-to-disk backups can be performed between two storage arrays or from a host to a storage array with the help of a host agent. As with other host-based data protection mechanisms, there may be specific requirements relative to operating system or application.

Disk-to-disk backup offers protection from most failures, such as various types of data center failures, and natural disasters. However, as with other backup methodologies, disk-to-disk backups are performed on a scheduled basis. In the event of a data loss failure at the primary site, any new or modified data since the last backup would be lost.

Recovery can involve restoration of specific files or entire file systems. Therefore recovery times can be anywhere from minutes to hours.

MIRRORING

File or block based

Mirroring can be file-based or block-based and can run at either the host level or the storage array level. File-based mirroring technologies track which directories and files have been modified so that the changes can be mirrored to the other location. This technology is also referred to as logical mirroring. Block-based mirroring tracks changed blocks and mirrors them to the other location. It typically runs at the storage array level and therefore is unaware of the specific directory and files being mirrored.

Host or storage array based

Host-based mirroring implements the copying of new and changed data using software agents that run on the hosts. This type of mirroring, although not necessarily dependent on specific storage hardware, has host operating system specific requirements. It also requires that the host have enough resources to run the mirroring software. It may also require specific mirroring methodologies for each application. Storage array-based mirroring reduces the mirroring overhead from the hosts and also has a single methodology regardless of the application or data being mirrored.

Asynchronous or synchronous

In an asynchronous mirror, changes are accumulated on the primary system and transmitted on a scheduled basis throughout the day. There are no delays to the applications because write operations are acknowledged immediately after the primary system receives them. Asynchronous mirroring typically does not have distance between the mirrors, but there is a tradeoff. If the primary system incurs a data-loss failure, any changes since the last mirror update are also lost. Asynchronous mirroring offers protection from regional disasters such as earthquakes and hurricanes.

In a synchronous mirror, once the application has generated a write operation, it cannot proceed until both the primary and the secondary have acknowledged it. The longer the secondary takes to acknowledge the operation, the longer the application is delayed. This is why the distance limitations are stricterwith synchronous mirroring. The advantage is that in the failure case described earlier, there would be no data loss. Synchronous mirroring offers complete data loss protection from localized disasters such as power and environmental failures.

Continuous asynchronous mirroring sends changes as they are being written rather than on a scheduled basis, but there are no delays to the applications because write operations are acknowledged immediately after the primary system receives them.

Recovery from a disaster typically involves a failover to the alternate site by making the mirror writable.

Table 1-4) RTO/RPO summary.

Method	RPO	RTO	Considerations
Point-in-Time Copies	Low	Low	Multiple recovery points
Physical Tape	High	High	Local or remote, offsite storage recommended
Disk-to-Disk Backup	Med	Med	Multiple recovery points in alternate locations
Asynchronous Mirror	Low	Low	Multiple locations, quick failover
Synchronous Mirror	0	Low	Zero data loss, limited distance

1.6 CHOOSING A DATA PROTECTION AND RECOVERY STRATEGY

GENERAL CONSIDERATIONS

At this point it is clear that there many choices of data protection methods. Which method is correct for your organization? The appropriate data protection strategy depends on an organization's specific requirements. A well-designed, scalable data protection strategy usually includes a combination of the methods discussed. This section discusses data protection in terms of tiers in order to provide examples of data protection strategies.

DATA CENTER PROTECTION

Table 1-5) Protection at the datacenter level.

Failure	Protection Method
Single disk	RAID4, RAID5, RAID6
Dual disk	RAID6
Multiple disk	RAID1 (local mirror)
Power supply and fan	Redundant components
Storage controller or processor	Redundant controllers
File deletions, data corruption, virus attacks, etc.	Point-in-time copies
	Physical tape
	Disk-to-disk backup
	Asynchronous mirroring
Power grid or environmental failure	Physical tape
	Disk-to-disk backup
	Asynchronous or synchronous mirroring

CAMPUS OR METROPOLITAN AREA PROTECTION

This level of protection offers all of the protection provided at the data center level plus the following.

Table 1-6) Protection at the campus or metropolitan area level.

Failure	Protection Method
Data center power outages	Mirroring , disk-to-disk backup, physical tape
Building fires	
Environmental (HVAC, plumbing)	

REGIONAL PROTECTION

This level of protection offers all of the protection provided at the data center and campus or metropolitan level plus the following.

Table 1-7) Protection at the regional level.

Failure	Protection Method
Electric grid failures	Asynchronous mirroring , disk-to-disk backup, physical tape
Natural disasters	
Terrorist incidents	

SAMPLE PROTECTION STRATEGIES

This section gives some examples of data protection strategies that combine one or more of the methods previously discussed. They are not all-inclusive and do not cover every conceivable combination; they are offered as an aid to assist in the development of the data protection plan that meets your organization's needs.

Point-in-timecopies plus physical tape backup (same data center)

Figure 1-4 shows an example of how data can be protected through a combination of point-in-time copies and the use of tape. The point-in-time copies can be run on a scheduled or a manual basis. This allows recovery to be performed up to the last copy made, and in some cases the user themselves are able to recover lost or deleted files instead of waiting for a system administrator. This reduces the recovery time and also labor costs associated with the administrator.

The tape referred to in Figure 1-4 could be physical, virtual, or a disk-to-disk backup system. It would normally be used for longer term protection and possibly archival purposes. It would also involve possibly greater recovery time and more labor for recovery.

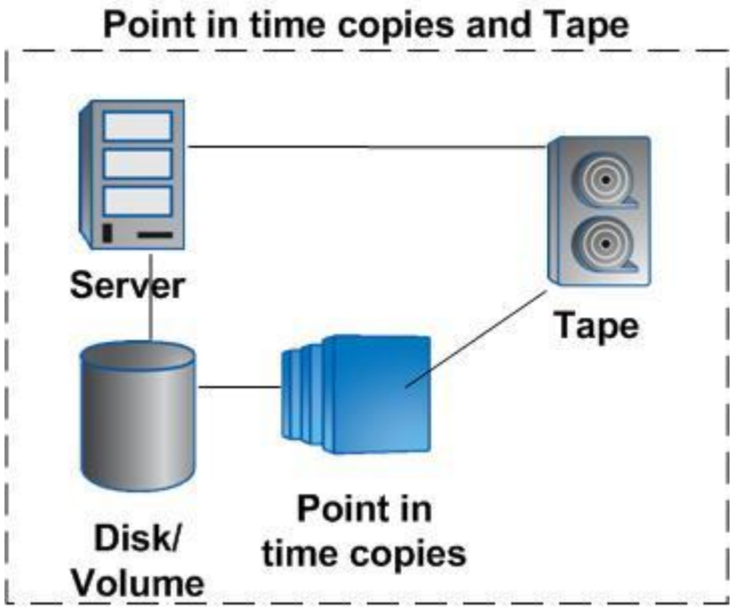


Figure 1-4) Point-in-time copies and tape.

Disk-to-disk plus physical tape (Campus)

Figure 1-5 shows a disk-to-disk backup solution that provides protection across locations (from Data Center 1 to Data Center 2). This strategy provides geographic protection (power, air conditioning, and so on) using disk-to-disk backup running on a scheduled basis. For longer term needs, the use of tape (or tape equivalent, as described earlier) is shown in Data Center 2, backing up the destination of the disk-to-disk backup. This offloads any additional burden from the production system in Data Center 1. Recovery from the disk-to-disk backup can be relatively fast compared to the slower archival tape device.

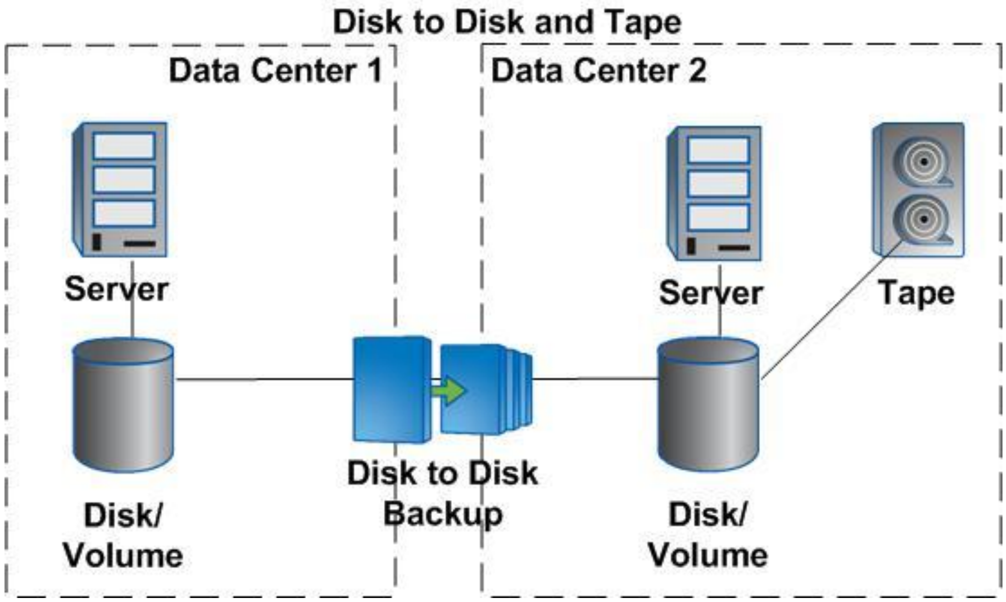


Figure 1-5) Disk-to-disk plus physical tape.

Asynchronous mirror plus disk-to-disk (regional)

Sometimes the recovery time and recovery point demands of the application (and consequently the data) require something more than periodic backup. In Figure 1-6, asynchronous mirroring is used to provide more frequent updates to the backup copy of the data. This reduces both the time required for recovery and the potential for data loss. Also shown is the use of disk-to-disk backupfrom the mirror instead of the production system. This reduces the load on the production system. Why use both methods? Mirroring is used to keep the destination as current as possible for minimal recovery point, but having backups with multiple recovery points can provide recovery back to any point in time (that is, protection from a virus attack).

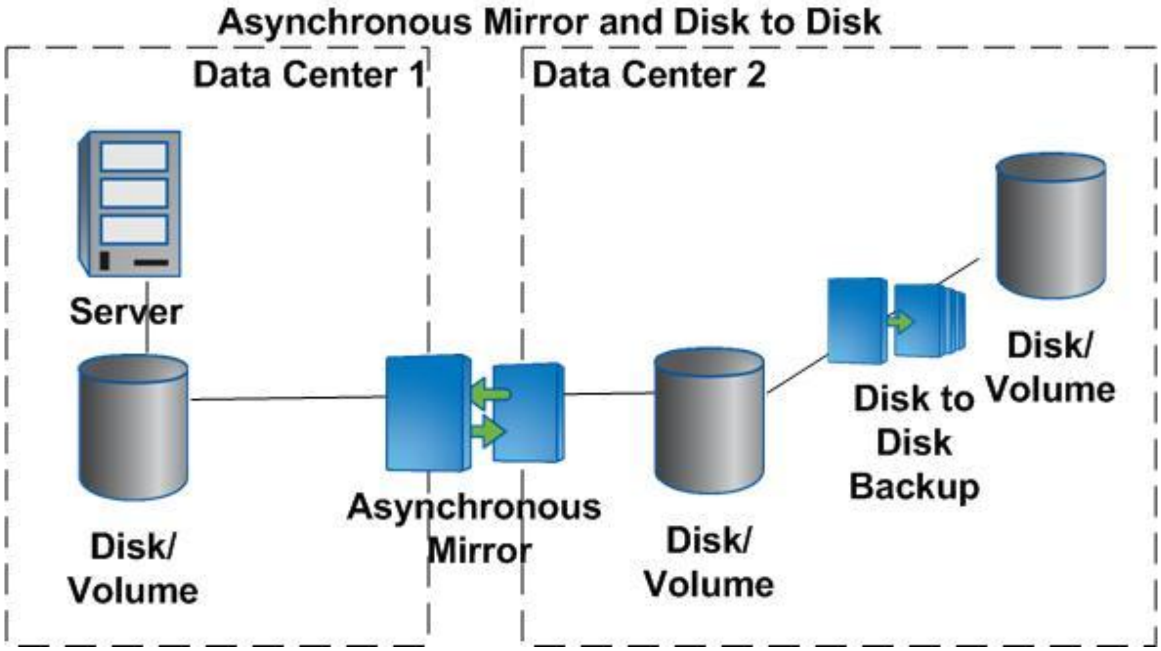


Figure 1-6) Asynchronous mirror and disk-to-disk.

Asynchronous mirror plus tape (regional)

Figure 1-7 is basically the same as Figure 1-6, except for the use of a tape library for archiving (or just for offsite disaster recovery).

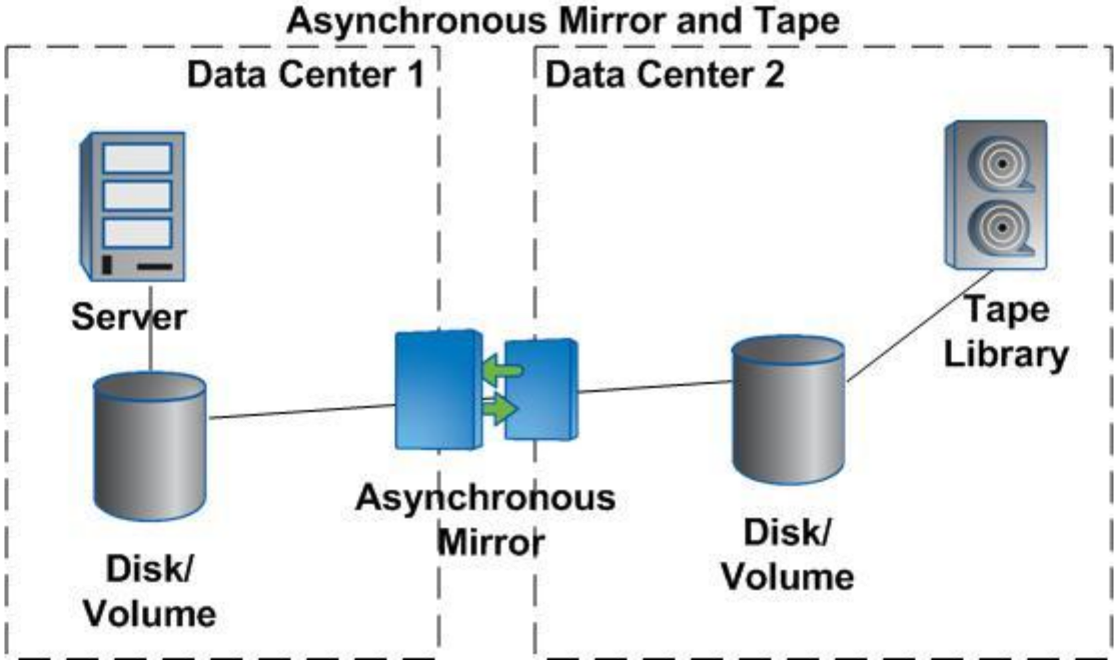


Figure 1-7) Asynchronous mirror and tape.

Synchronous mirror plusdisk-to-disk (data center, campus, regional)

At times, the criticality of the application and data demands zero data loss protection and a minimal recovery time, making synchronous mirroring a requirement. As shown in Figure 1-8, synchronous mirroring is used between locations to meet that need. However, because of the synchronous nature, there are limitations on the distance between locations. Similar to Figure 1-6, in this case disk-to-disk backup is performed at the alternate location to provide multiple recovery points without additional burden to the production system.

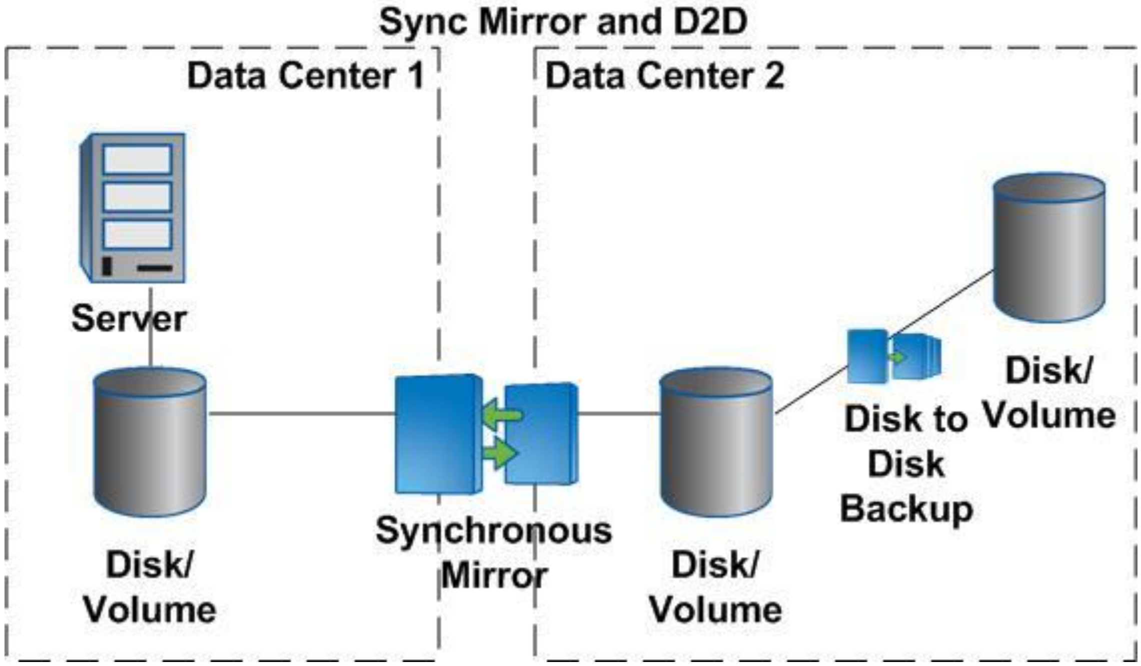


Figure1-8) Synchronous mirror and disk-to-disk.

Synchronous mirror, asynchronous mirror (DC, campus, regional)

Figure 1-9 shows a multi-tier protection strategy that uses synchronous mirroring to provide continuous data availability between data centers and asynchronous mirroring for longer distance protection for disaster recovery.

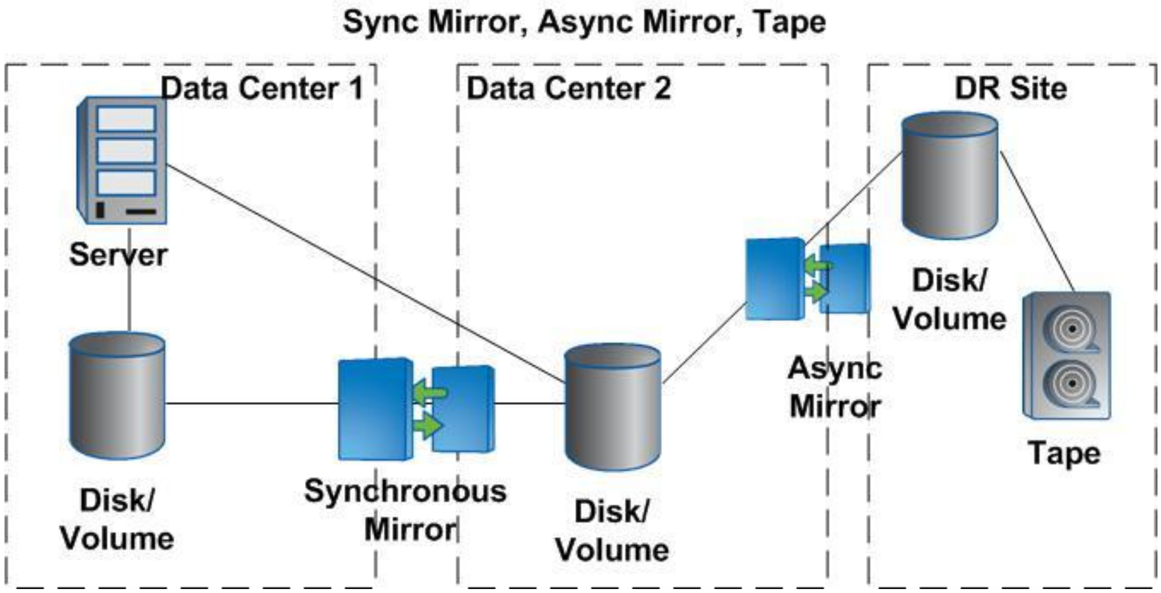


Figure 1-9) Point-in-time copies and tape.

1.7 GLOSSARY

applicationconsistency. The state in which all related data components (databases, flatfiles, and so on) are in a transaction-consistent state and are in synchronization based on the application design and requirements.

applicationrecovery. The component of disaster recovery that deals specifically with the restoration of business system software and data after the processing platform has been restored or replaced.

backup.A process by which electronic or paper-based data is copied in some form so that it is available if the original data is lost, destroyed, or corrupted.

businesscontinuity. The ability of an organization to provide service and support for its customers and to maintain its viability before, during, and after a business continuity event, such as data loss due to human error or a natural disaster.

businesscontinuity planning. The process of developing and documenting arrangements and procedures that enable an organization to respond to an event that lasts for an unacceptable period of time and return to performing its critical functions after an interruption.

businesscontinuity strategy: An approach by an organization to ensure its recovery and continuity in case of a disaster or other major outage. Plans and methodologies are determined by the organization’s strategy; there may be more than one solution. **Examples:** Internal or external hotsite or coldsite;alternate work area reciprocal agreement;mobile recovery, quick ship and drop ship;consortium-based solutions; etc.

businessimpact analysis:A process designed to prioritize business functions by assessing the potential quantitative (financial) and qualitative (nonfinancial) impacts that could result if an organization experiences a business continuity event, such as data loss due to human error or a natural disaster.

businessinterruption: Any event, whether anticipated (for example, a public service strike) or unanticipated (for example, a blackout) thatdisrupts an organization’snormal course of business operations.

coldsite: An alternate facility that already has in place the environmental infrastructure required to recover critical business functions or information systems, but does not have any preinstalled computer hardware, telecommunications equipment, communication lines, etc. These must be provisioned at the time of a disaster.

continuity of operations plan:A COOP provides guidance on system restoration for emergencies, disasters, and mobilization, and for maintaining a state of readiness to provide the necessary level of information-processing support commensurate with the mission requirements and priorities identified by the respective functional representative. The federal government and its supporting agencies traditionally use this term to describe activities otherwise known as disaster recovery, business continuity, business resumption, or contingency planning.

continuousavailability:-A system or application that supports operations thatcontinue with little orno noticeable impact onthe user in the event of a failure such as a power outage. For instance, with continuous availability, the user doesnot have to log in again or to resubmit a partial or whole transaction.

continuousoperation: The ability of an organization to perform its processes without interruption.

criticalbusiness functions:The critical operational and/or business support functions that could not be interrupted or unavailable for more than a mandated or predetermined time without significantly jeopardizing the organization. An example of a business function is a logical grouping of processes or activities that produce a product and/or service, such as accounting, staffing, customer service, etc.

criticaldata point: The point in time to which data must be restored in order to achieve recovery objectives.

databackups: The copying of production files to media that can be stored on and/or offsite and can be used to restore corrupted or lost data or to recover entire systems and databases in the event of a disaster.

databackup strategies: Data backup strategies determine the technologies, media, and offsite storage of the backups necessary to meet an organization’s data recovery and restoration objectives.

datacenter recovery: The component of disaster recovery thatdeals with the restoration of data center services and computer processing capabilities at an alternate location and the migration back to the production site.

datamirroring:A process whereby critical data is replicated to another device.

dataprotection: The process of ensuring the confidentiality, integrity, and availability of data

datarecovery: The restoration of computer files from backup media to restore programs and production data to the state that existed at the time of the last safe backup.

databasereplication: The partial or full duplication of data from a source database to one or more destination databases.

disaster:A sudden, unplanned catastrophic event causing unacceptable damage or loss. 1) An event that compromises an organization’s ability to provide critical functions, processes, or services for some unacceptable period of time. 2) An event where an organization’s management invokes their recovery plans.

disasterrecovery: The ability of an organization to respond to a disaster or an interruption in services by implementing a disaster recovery plan to stabilize and restore the organization’s critical functions.

disasterrecovery plan: The management-approved document that defines the resources, actions, tasks, and data required to manage the technology recovery effort. This is a component of the business continuity plan.

disasterrecovery planning: The technical component of business continuity planning.

disk-to-disk backup: With D2D, a computer hard disk is backed up to another hard disk rather than to a tape or floppy. D2D is often confused with virtual tape, but differs in that it enables multiple backup and recovery operations to simultaneously access the disk directly by using a true file system.

event: Any occurrence that may lead to a business continuity incident.

gapanalysis:A detailed examination to identify risks associated with the differences between business and operations requirements and the current available recovery capabilities.

highavailability: Systems or applications that requirea very high level of reliability and availability. High availability systems typically operate 24/7 and usually require built-in redundancy to minimize the risk of downtime due to hardware and/or telecommunication failures.

hotsite: An alternate facility that already has in place the computer, telecommunications, and environmental infrastructure required to recover critical business functions or information systems.

impact: The effect, acceptable or unacceptable, of an event on an organization. The types of business impact are usually described as financial and nonfinancial and are further divided into specific types of impact.

latency:(1)The amount of time between the time when a network device receives a packet and the time when the packet is retransmitted.

(2) The amount of time between the instant at which an instruction control unit initiates a call for data and the instant at which the actual transfer of the data starts.

(3) The time from the initiation of an operation until something actually starts happening (for example, data transmission begins).

(4) In replication, part or all of the approximate difference between the time that a source table is changed and the time that the change is applied to the corresponding target table.

mirroring:(1) The process of writing the same data to two disk units inthe same auxiliary storage pool at the same time. The two disk units become a mirrored pair, allowing the system to continue when one of the mirrored units fails.

(2) The process of writing the same data to multiple disks at the same time. The mirroring of data protects it against data loss inthe database or inthe recovery log.**mission-critical applications:**Applications that support business activities or processes that could not be interrupted or unavailable for 24 hours or less without significantly jeopardizing the organization.

offsite storage: Any place physically located a significant distance away from the primary site, where duplicated and vital records (hard copy or electronic and/or equipment) canbe stored for use during recovery.

outage: The interruption of automated processing systems, infrastructure, support services, or essential business operations, which may result, in the organization’s inability to provide services for some period of time.

risk: Potential for exposure to loss, which can be determined by using either qualitative or quantitative measures.

recovery: Implementing the prioritized actions required to return the processes and support functions to operational stability following an interruption or disaster.

recoverypoint objective: The maximum amount of data loss an organization can sustain during an event, as defined by the organization.

recoverysite:A designated site for the recovery of business unit, technology, or other operationsthatare critical to the enterprise.

recoverytime objective: The period of time within which systems, applications, or functions must be recovered after an outage, as defined by the organization (for example, one business day). RTOs are often used as the basis for the development of recovery strategies and to determine whether or not to implement the recovery strategies during a disaster situation.

replication:(1) The process of maintaining a defined set of data in more than one location. Replication involves copying designated changes for one location (a source) to another (a target) and synchronizing the data in both locations.

(2) The process of copying objects from one node in a cluster to one or more other nodes in the cluster, which makes the objects on all the systems identical.

(3) The process of exchanging modifications between replicas, making all of the replicas essentially identical over time.

resilience: The ability of an organization to absorb the impact of a business interruption and continue to provide a minimum acceptable level of service.

service-level agreement:A formal agreement between a service provider (whether internal or external) and their client (whether internal or external), which covers the nature, quality, availability, scope, and response of the service provider. The SLA should cover both day-to-day situations and disaster situations, becausethe need for the service may vary in a disaster.

singlepoint of failure:A unique pathway or source of a service, activity, and/or process. Typically there is no alternative, and a loss of that element could lead to a failure of a critical function.

systemrecovery: The procedures for rebuilding a computer system and network to the condition where it is ready to accept data and applications and to facilitate network communications.

virtualtapelibrary:A data storagevirtualization technology used typically for backup and recovery. A VTL presents a storage component (usually hard disk storage) as tape libraries or tape drives for use with existing backup software.

warmsite: An alternate processing site thatis equipped with some hardware, as well as communications interfaces and electrical and environmental conditioning thatis capable of providing backup only after additional provisioning customization is performed or software is added.

2 NETAPP DATA PROTECTION SOLUTIONS

2.1 INTRODUCTION

Volume I of the Data Protection Handbook introduced and defined challenges that threaten the availability and recoverability of critical data. In addition, it outlined some of the data protection strategies used in the industry today. The next step is to understand how to implement these strategies.

Volume II introduces the various NetApp products and solutions available to address the many data protection challenges and strategies. It is important to understand the characteristics of each product as well as how they work together. This volume covers NetApp solutions for backups, mirroring, compliance, application consistency, and storage efficiency, as well as our management application.

Before discussing the individual products,it’s important to understand NetApp Snapshot™ copies.

NETAPP SNAPSHOT COPIES

NetApp storage systems include the core capability to create instant read-only point-in-time copies of data, called Snapshot copies. Snapshot copies are primarily used as local backups for rapid recovery. However, several data protection products also take advantage of these copies in order to replicate block changes to remote storage.

Snapshot copies are user accessible, and each copy appears as a full backup of the file system even though the space consumed represents only blocks that have changed since the previous Snapshot copy was taken.

2.2 NETAPP DATA PROTECTION AND RECOVERY METHODS

This section describes the data protection and recovery methods offered by NetApp.

NETAPP BACKUP

Tape

Although tape backup technologies have their challenges, it is still a common requirement for many enterprises to incorporate tape into the data protection solution. This is especially true when backups need to be stored offsite. Quite often tape is used to complement more advanced methods of data protection and to satisfy regulatory requirements. Tape backups from NetApp storage systems can be accomplished with most backup applications that take advantage of NDMP standards. However, recovery point objectives (RPO) and recovery time objectives (RTO) associated with tape often make it a less than desirable strategy.

SnapVault

SnapVault® is a storage-based disk-to-disk backup and recovery solution that protects NetApp primary storage. It leverages the efficiencies of Snapshot copies and protects data at the block level. This is fundamentally different than legacy tape-based approaches where backups are done on a file level. This block-level solution requires only a single full backup. Once the initial full backup is complete, only changed blocks are replicated to the NetApp secondary system. This is a true “incremental forever” technology.

Figure 2-1 illustrates SnapVault protecting NetApp storage.

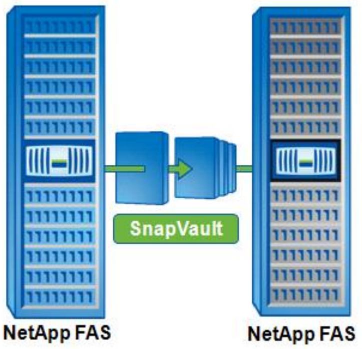


Figure 2-1) SnapVault protects NetApp storage.

The efficiencies built into SnapVault make it possible to drastically reduce backup and recovery times, allowing tighter RPO and RTO. SnapVault backups also work well over WANs where NetApp primary storage is deployed in remote offices. Additionally, the capacity requirements for long-term retention are much lower compared to tape-based solutions. And when combined with deduplication (see "Deduplication," later in this section), these long-term requirements can be even less.

With SnapVault, recoveries can include individual files and directories as well as entire qtrees. When restoring an entire qtree, only those blocks required to recover the qtree to a particular point in time are restored.

For more information about SnapVault, see the *SnapVault Best Practices Guide*.

Open Systems SnapVault

Open Systems SnapVault is an agent-based extension of SnapVault thatoffers a heterogeneous block-level backup solution for open systems platforms. Similar to SnapVault, Open Systems SnapVault works at the block level, allowing efficient backups over WANs.

With Open Systems SnapVault, file systems from Windows®, Linux®, and UNIX®can be protected and retained on NetApp secondary storage. These heterogeneous backups are stored in native format and can be directly accessed for drag-and-drop restores. In addition, VMware® virtual machines that reside on third-party storage can be protected with the ESX agent, allowing simplified recovery of entire virtual machines.

Figure 2-2 shows an example of Open Systems SnapVault.

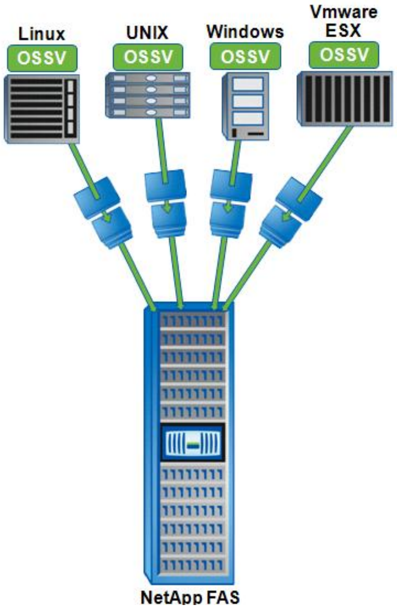


Figure 2-2) Open Systems SnapVault.

With Open Systems SnapVault, recoveries can include individual files and directories as well as entire file systems. When protecting VMware virtual machines, recoveries are limited to the entire virtual machine.

For more information aboutOpen Systems SnapVault, see the *Open Systems SnapVault Best Practices Guide* and the *Open Systems SnapVault Best Practices Guide for Protecting Virtual Infrastructure*.

NETAPP MIRRORING

Data mirroring is significant in designing the overall data protection solution. When building disaster recovery into your data protection scheme, consider the following NetApp products.

SnapMirror

SnapMirror® is a data mirroring product that offers lower RPO than backup and recovery solutions. In addition, SnapMirror can easily fail over to the secondary site, allowing low RTO. Once the primary site is available again, services can be failed back for normal operation. SnapMirror works at the block level and produces an exact copy of the data at the remote site. Three modes of replication are available with SnapMirror: asynchronous, synchronous, and semi synchronous.

Figure 2-3 shows SnapMirror protecting NetApp storage.

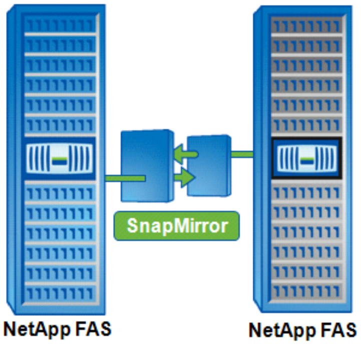


Figure 2-3) SnapMirror protects NetApp storage.

With SnapMirror Async, changed blocks captured by Snapshot copies are replicated to the destination in intervals as frequently as once per minute. SnapMirror Async is an extremely efficient form of replication. In addition to replicating at the block level, it can also replicate in a deduplicated state (see ”Deduplication,” later in this section). For more information about SnapMirror Async, see the *SnapMirror Async Overview and Best Practices Guide*.

SnapMirror Sync replicates continuously as each write occurs on the primary, enabling zero data loss. The cost, however, is write performance. As with most synchronous replication, writes aren’t complete until they reach the mirror site.For more information about SnapMirror Sync, see the technical report *SnapMirror Sync and SnapMirror Semi-Sync Overview and Design Considerations*.

SnapMirror Semi-Sync is similar to SnapMirror Sync. However, with semi-synchronous replication, writes are acknowledged immediately. This mode of replication provides a low RPO while maintaining write performance. For more information aboutSnapMirror Semi-Sync, see the technical report*SnapMirror Sync and SnapMirror Semi-Sync Overview and Design Considerations*.

MetroCluster

As shown in Figure 2-4, MetroCluster provides high-performance synchronous mirroring over Fibre Channel as well as continuous availability between locations up to 100 kilometers apart.MetroCluster uses controller-based mirroring and can automatically recover from single component failures. For site outages, MetroCluster provides single-command failover capabilities for disaster recovery. The value of MetroCluster is that it minimizes downtime with no data loss.

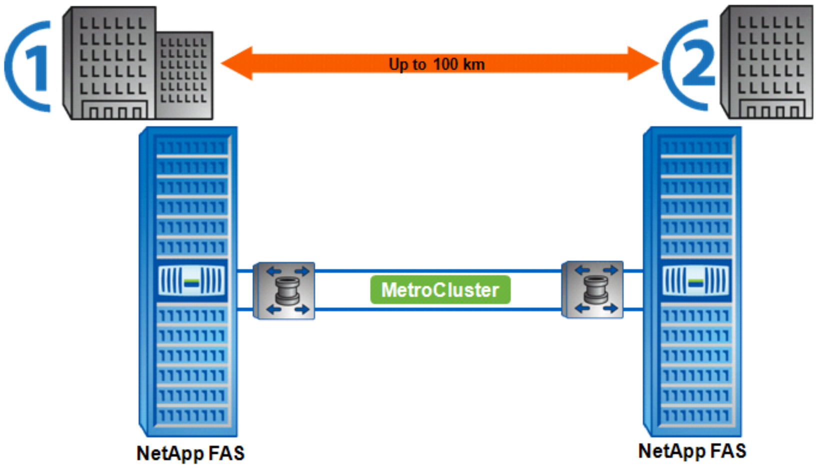


Figure 2-4) MetroCluster.

For more information about MetroCluster, see the *MetroCluster Design and Implementation Guide*.

NETAPP COMPLIANCE

Many organizations today are burdened with regulatory mandates regarding the preservation of data. Not only must records be kept for years, they must be unalterable and unerasable. SnapLock® provides secure data permanence on NetApp storage and can be further protected when combined with backup and mirroring solutions such as SnapVault, SnapMirror, and MetroCluster.

SnapLock compliance

SnapLock Compliance adheres to government and industry standards such as SEC Rule 17a-4, HIPAA, Sarbanes-Oxley, 21CFR Part 11, and DOD 5015.2. With SnapLock Compliance, data can never be altered or removed before the retention date expires. Further, the compliant volume can never be destroyed short of willful destruction such as physically removing disks from the system.

SnapLock enterprise

SnapLock Enterprise is very similar to SnapLock Compliance except that it adheres to best practices set forth by the organization. With SnapLock Enterprise, data remains unalterable and unerasable, but an administrator can delete an entire SnapLock Enterprise volume if necessary.

For more information aboutNetApp compliance, see the SnapLock Compliance and SnapLock Enterprise Software website.

APPLICATION CONSISTENCY

NetApp's high-performance primary storage is well suited for application and database data. The resiliency of RAID-DP® technology built into every NetApp system offers a high tolerance to disk failures. However, data loss due to operational mistakes and corruption is arisks that must be addressed. In addition, application data must be in a consistent state before it can be backed up. When application data is unavailable, restoring to a known good state as quickly as possible becomes an urgent matter. Having frequent backups local to the primary system means speedy recoveries and reduced data loss when restores are needed.

Several SnapManager® products are available to accomplish local backups and restores for application data. SnapManager works with the application to create a consistent Snapshot copy. Because Snapshot copies capture only changed blocks, storage requirements for retention can be quite low. SnapManager is integrated with the following applications: Microsoft® Exchange, Oracle®, Microsoft SQL Server®, Microsoft Office SharePoint® Server, VMware, and SAP®.

For more information aboutthese SnapManager products:

SnapManager for Microsoft Exchange

SnapManager for Oracle

SnapManager for Microsoft SQL Server

SnapManager for Microsoft Office SharePoint Server

SnapManager for Virtual Infrastructure

SnapManager for SAP

LEVERAGING STORAGE EFFICIENCIES

Deduplication

NetApp deduplication has a fundamental advantage over many deduplication technologies on the market today. Althoughmany vendors offer deduplication only on secondary storage for backup destinations, NetApp deduplication can run on primary or secondary storage.

NetApp deduplication running on primary storage offers significant space saving potential for user data as well as critical application data. NetApp deduplication has shown extreme space savings in VMware environments where virtual machine images often contain much of the same data.

Both SnapMirror and SnapVault are integrated with NetApp deduplication. As shown in Figure 2-5, SnapMirror replicates blocks in a deduplicated state, reducing the amount of data sent over the network during transfers and reducing storage requirements at the mirror site.

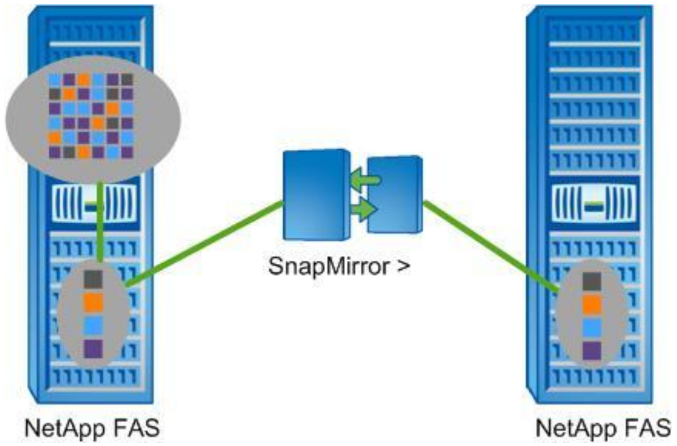


Figure2-5) SnapMirror replication of deduplicated blocks.

Integration with SnapVault means that deduplication on the secondary, when enabled, runs automatically after each transfer. This reduces storage requirements for long-term backup retention. In Figure 2-6, multiple backups are directed to a single destination volume, where deduplication is performed against all of the backup data. The same approach also works with Open Systems SnapVault backups.

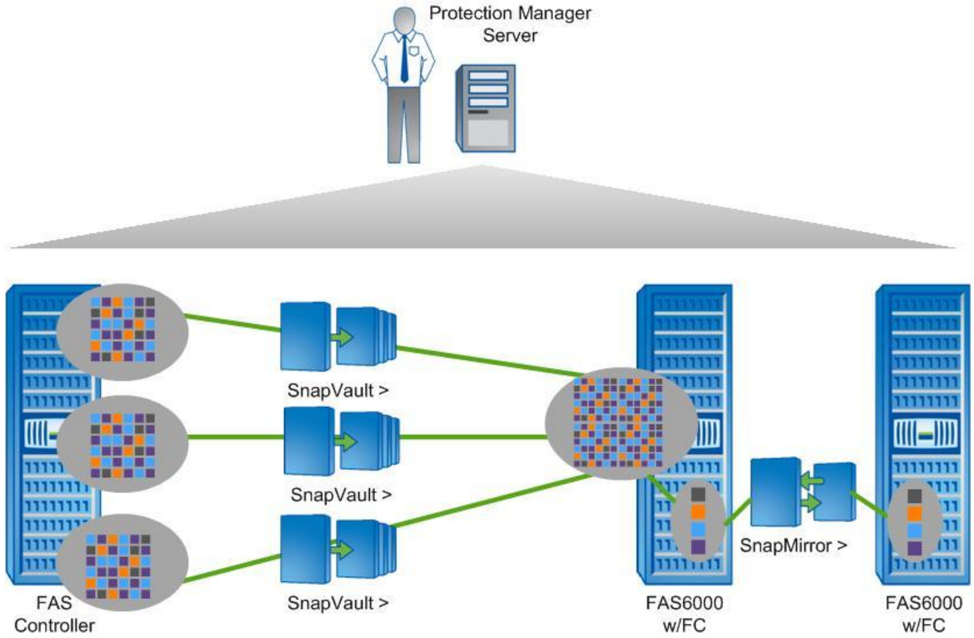


Figure 2-6) Deduplication of SnapVault backups

For more information, seeNetApp Deduplication. In addition, you can use the NetApp deduplication calculator to get an idea of the savings you can expect.

FlexClone

NetApp FlexClone® technology enables true cloning—instant replication of data volumes, files, and LUNs without requiring additional storage space at the time of creation. Each cloned LUN or volume is a transparent, virtual copy of your data 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 provisioning of physical and/or virtual server and desktop images. FlexClone provides substantial storage space savings, which means that you can manage many more data set variations in less time and with less risk. In addition to storage savings, FlexClone can help reduce power and cooling costs through efficient storage utilization.

For example, ifa FlexClone copy is created from a SnapVault backup, it can be used for reporting.In addition, you can be sure that theSnapVault backups are successful becauseyou can bring up the database by using this FlexClone copy. A FlexClone copy can also be taken from a SnapMirror destination, as shown in Figure 2-7. When created from a SnapVault backup or SnapMirror destination, using a FlexClone copy has no impact to the production system.

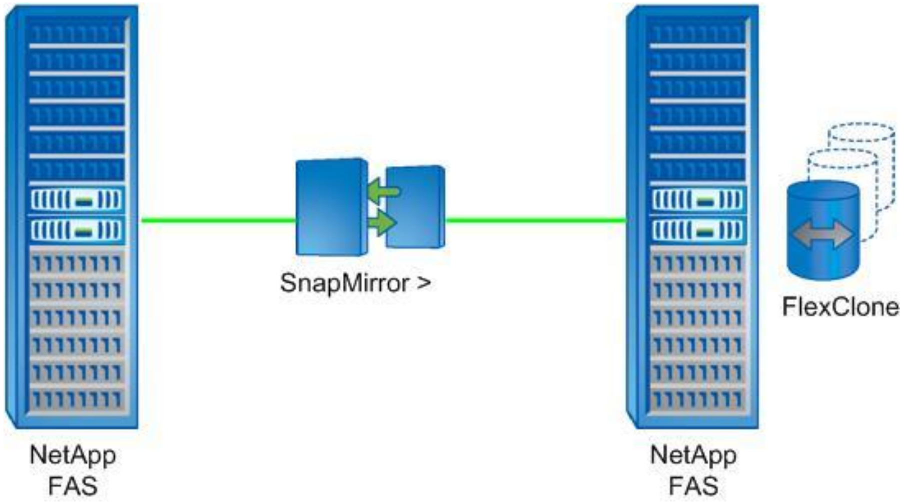


Figure 2-7) FlexClone copies of SnapMirror destination data

For development and testing, imagine a single source of engineering code. With multiple FlexClone copies, developers can modify and test their copy with no risk to the original source. When finished, the FlexClone can be destroyed or splitoff and used as the new source data.

MANAGEMENT APPLICATION

Data protection solutions can get quitecomplex. Management applications alleviate some of these challenges by simplifying routine tasks, scheduling conflicts, and recoveries. Protection Manager is the NetApp management console that does just that.

Protection Manager

Protection Manager is a policy-based application for managing disk-to-disk backup and recovery as well as DR replication. Protection Manager policies specify what technology is used to protect the data, eliminating much of the guesswork. Protection Manager can automatically provision and expand the required destination storage as needed.

- Allows global monitoring of all data protection tasks and storage resources
- Enables rapid protection of new systems and discovers unprotected data
- Identifies issues and recommends solutions
- Increases utilization of storage resources

Figure 2-8 shows Protection Manager facilitating backup and mirroring operations.

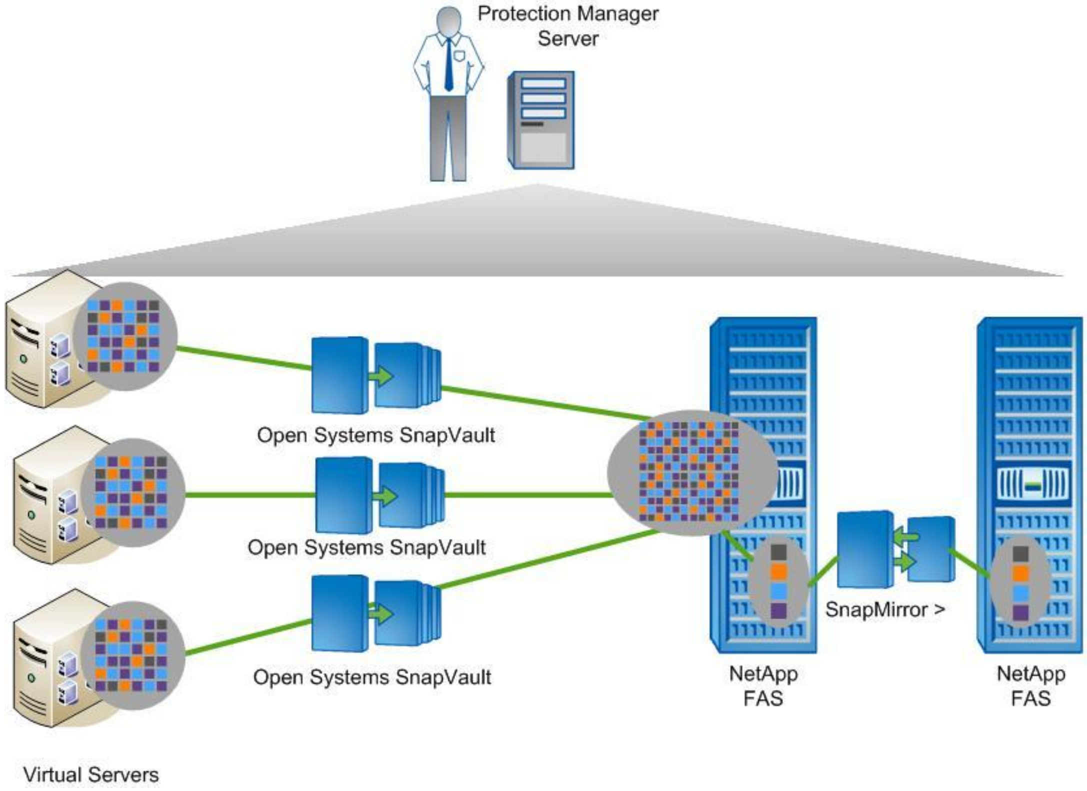


Figure 2-8) Managing backups and mirroring with Protection Manager

Volume III of thisData Protection Handbook provides examples of how Protection Manager facilitates different data protection scenarios.

For detailed information aboutProtection Manager, see the *Protection Manager Best Practices Guide*.

2.3 CHOOSING THE RIGHT NETAPP SOLUTION

This section provides two examples, combining several data protection methods to form different solutions. Figure 2-9 illustrates a solution weighted heavily toward backup and recovery. In this approach, SnapVault and Open Systems SnapVault are used in each datacenter to accomplish disk-to-disk backup and recovery. The backup targets in each datacenter mirror each other, providing offsite DR protection of the backup data. The benefit here is that there are redundant copies of the backup data, and those copies are geographically separated. In addition, if something should happen to the SnapVault destination system, backups can be redirected to the remote site.

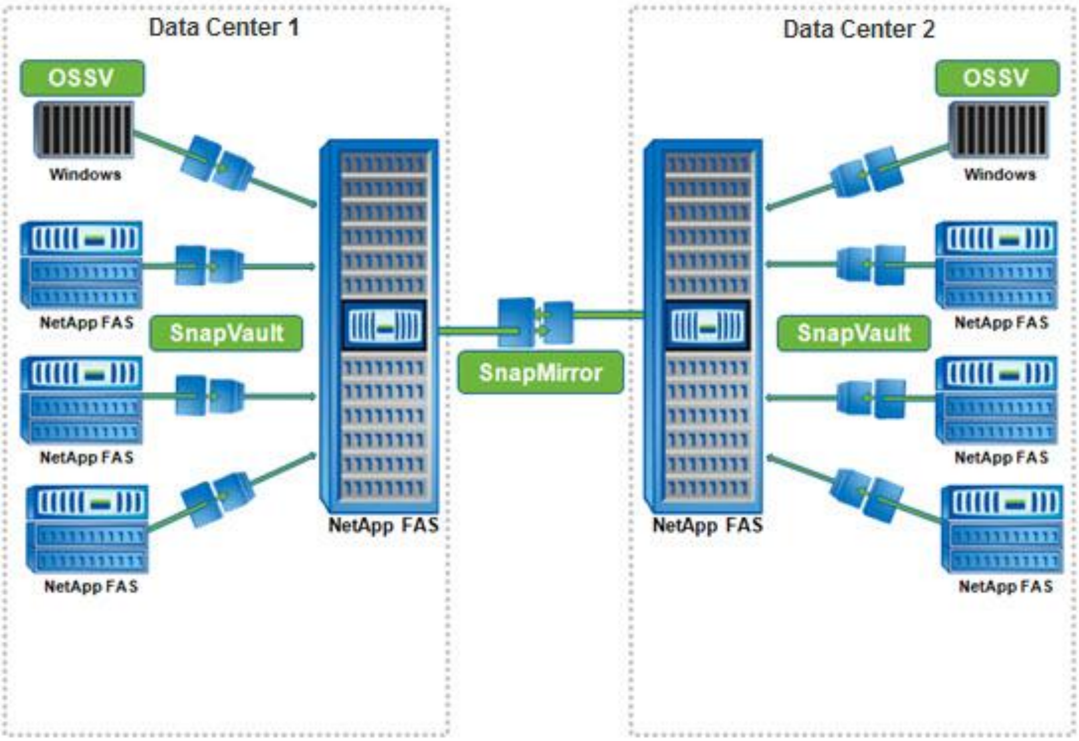


Figure 2-9) Local backups mirrored to remote location.

In Figure 2-10, Data Center 1 is the primary site. SnapMirror provides a DR strategy for the primary data. Backups are taken from the DR system in Data Center 2, enabling offsitebackups. Backup processes are done in the remote site, freeing up cycles on the primary system. Additionally, backups to tape can be generated from the DR site if tape is a requirement.If Data Center 1 becomes unavailable, services can fail over to Data Center 2.

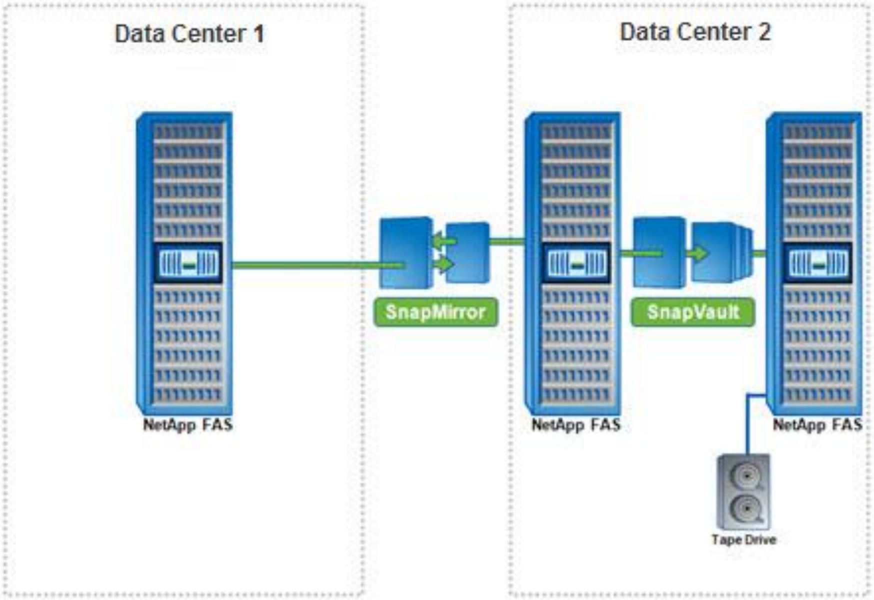


Figure. 2-10) Backups taken from mirror at remote location

2.4 SUMMARY

There are many challenges for enterprises to overcome regarding data protection. However, most organizations don’t have unlimited budgets to address these challenges. NetApp has a variety of products tomeet these data protection needs in a cost-effective and space-efficient manner.

Obviously, there is no single product than can solve all business problems. In most cases, multiple products are combined to form a total solution where the individual products complement each other regardless of their interoperability characteristics. For example, NetApp MetroCluster and SnapVault can complement each other as part of a complete enterprise solution by providing campus availability and D2D backups for long term retention. Understanding the key values of each product makes it easier to develop solutions that address each identified challenge.

3 EXAMPLE SCENARIOS

This section walks through some scenarios and highlights how NetApp products can offer a comprehensive data protection solution with various configurations. These scenarios detail specific considerations and steps to be followed when implementing the data protection solutions. Best practices for the scenarios are also discussed. For general installation and configuration information,see the product documentation.

3.1 SCENARIO 1: DISK-TO-DISK BACKUP AND ASYNCHRONOUS MIRRORING

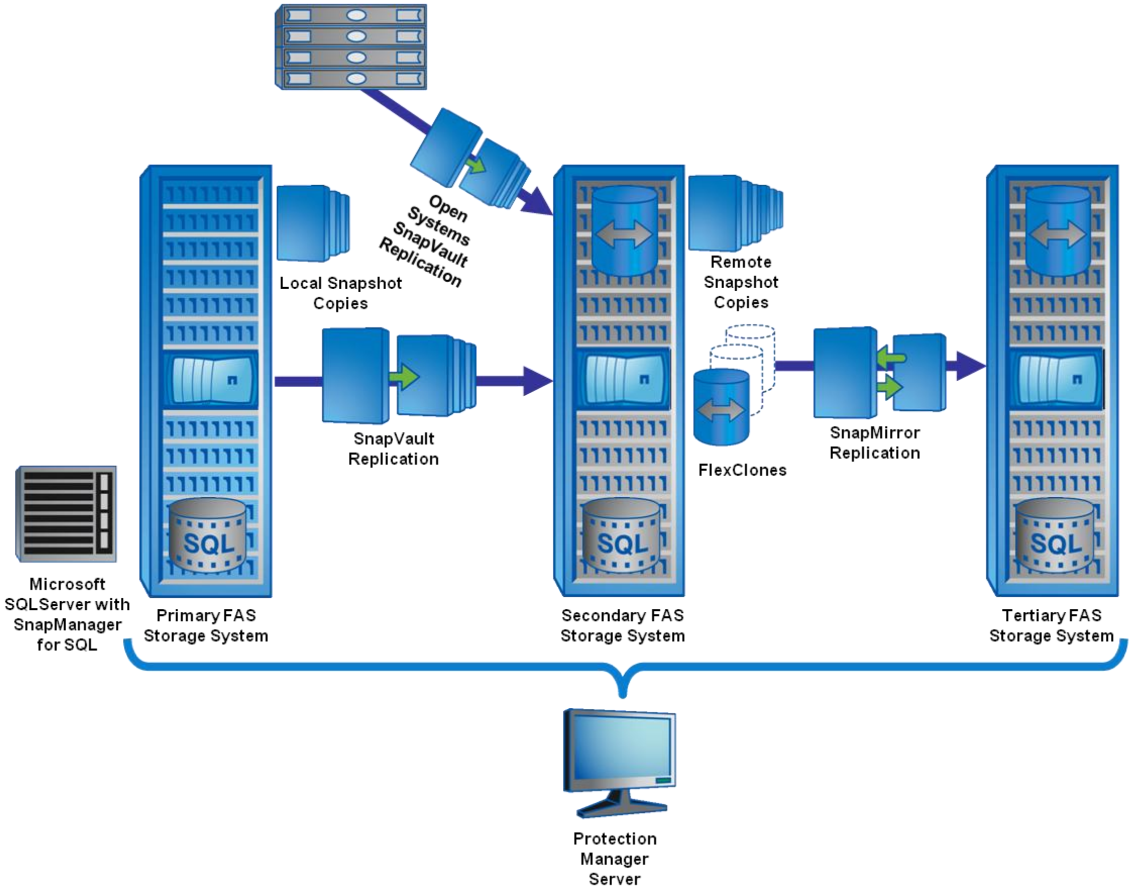

Open Systems SnapVault Replication
Local Snapshot Copies
Remote Snapshot Copies
SnapVault Replication
FlexClones
SnapMirror Replication
SQL
SQL
SQL
Microsoft SQLServer with SnapManager for SQL
Primary FAS Storage System
Secondary FAS Storage System
Tertiary FAS Storage System
Protection Manager Server

Figure 3-1) Scenario 1 environment.

OVERVIEW

This scenario shows how NetApp data protection and application integration products integrate to provide an efficient and comprehensive disk-to-disk backup solution. NetApp SnapManager® for SQL Server® and data replication technologies offer fast and space-efficient backup copies for both SQL databases and remote office data with a single management interface with Protection Manager. The following sections in this scenario describe and illustrate how the overall solution is configured and some of the steps implementing such a solution.

BUSINESS PROBLEMS AND PROTECTION PROVIDED

This scenario can help solve multiple business problems, starting with the reduction or elimination of tape in the environment. In addition, multiple recovery points can be stored at various locations, providing a robust backup and recovery solution.

In this environment (see figure 3-1), there are three tiers of protection with many recovery points and options for recovery location. In case of local corruption, there is the option of recovery from local snapshots (this is typically the first option and would minimize RTO). Given that these recovery points reside on more expensive FC disks, only limited recovery points are kept locally. The next option is to use a snapshot copy from the secondary storage system. With the addition of replicating the backups to another system, the secondary system is also protected, providing another repository of recovery points. Typically, the secondary and tertiary storage systems have larger, cheaper SATA disk drives, so storing historical data on these systems has a lower TCO than storing the data on the primary storage system.

IMPLEMENTATION

The configuration of this scenario uses multiple NetApp technologies and software products. This section describes the environment in this scenario and then dives into the details of the implementation. As shown in figure 3-1, NetApp Protection Manager has been installed to manage the data replication and monitor the environment. Protection Manager allows easy and simple policy-based management of this environment. In this scenario, Protection Manager is used in to monitor and manage the local backup replication (SnapVault or Open Systems SnapVault) of the data and then the offsite replication of the data (volume SnapMirror).

There is also a Microsoft SQL Server running on iSCSI LUNs presented from a NetApp storage system. SnapManager for SQL Server has been installed on the SQL Server to help automate backups for the SQL database. SnapManager for SQL Server in this scenario handles the creation and retention of SQL-consistent Snapshot copies on the local NetApp system. With the addition of Protection Manager to this environment, SnapManager for SQL Server and Protection Manager integrate to aid in the replication of the SQL backup data offsite, using SnapVault. Protection Manager maintains the retention of the SQL backups on the SnapVault destination. As part of the SnapManager for SQL Server backup process, there is the option for verification of the backup data. Adding SnapVault to the solution gives the ability to perform the backup verification from the remote site if a SQL Server is present at that site.

In addition to SQL Server, some systems run Windows 2003 in a remote office with limited bandwidth back to the data center. By installing Open Systems SnapVault on these remote servers, backup data can be centralized on the same system as the SQL offsite backup data. Open Systems SnapVault allows the data to be efficiently transferred over the wire by sending only the 4K blocks that have changed, and storing only those blocks on disk.

To satisfy the offsite requirement, all backup data is then replicated, once a day, to an alternate site using volume SnapMirror. This allows all backup data to be stored outside of the centralized data center and provides another copy of the data that can be used in the event of a disaster. In addition, if there is an issue with the centralized backup storage system, backup replication can continue to the volume SnapMirror destination system without the need to perform a baseline transfer.

A common theme across the entire implementation is the value of reducing the amount of data on disk, and the amount of bandwidth required to replicate that data. As of Data ONTAP® 7.3, SnapVault and Open Systems SnapVault are integrated with NetApp deduplication on FAS. If a SnapVault destination volume has deduplication enabled, after each SnapVault transfer is complete, the duplicate blocks in the last incremental update are deduplicated and the free space returned to the volume. This functionality is also integrated into Provisioning Manager to allow the provisioning policy to create a volume with deduplication turned on after the volume is created.

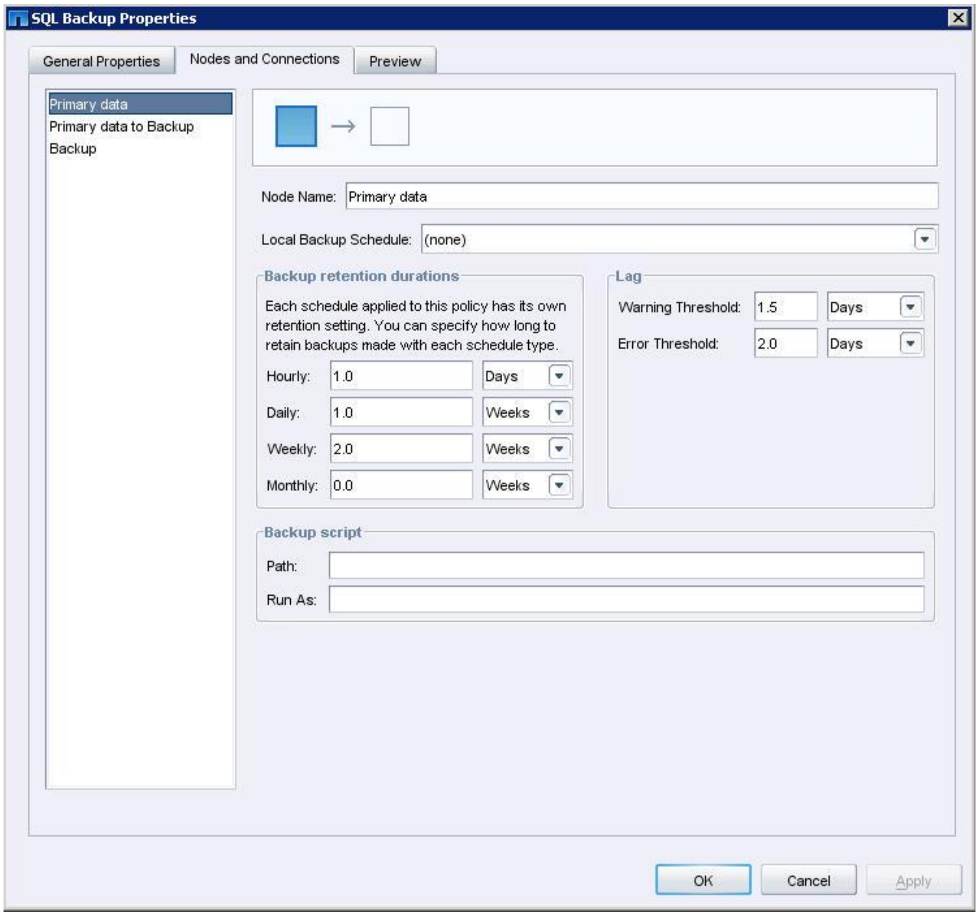
Note: Once deduplication is turned on for the destination volume, scheduled or manual deduplication processes are disabled.

Now that all the components have been described at a high level, let’s look more closely at the architecture and implementation that are specific to this scenario. For more information about the general installation and configuration of these products, refer to the appropriate product documentation.

SnapManager for SQL server: installation and configuration

After the required products (SnapDrive, SnapManager for SQL Server,and Protection Manager) have been installed, only a few steps are required to bring all the pieces together. The first step is to create a valid Protection Policy within Protection Manager. There are some restrictions that must be followed with regard to the Protection Policy for the SnapManager for SQLServer integration to function properly. In the following example, the “Remote backup only” Protection Policy was copied and named SQL Backup. After the Protection Policy has been created, the following configuration changes to it are required.

1. The local (primary data) backup schedule must be set to None(Figure 3-2).   In this environment, SnapManager for SQL Server controls the local Snapshot creation and retention. Allowing SnapManager for SQL Server to control the creation and deletion of the Snapshot copies means that the Snapshot is a valid and consistent copy of the SQL database.



Primary data
Primary data to Backup
Backup

Node Name: Primary data

Local Backup Schedule: (none)

Backup retention durations

Each schedule applied to this policy has its own retention setting. You can specify how long to retain backups made with each schedule type.

Hourly:	1.0	Days
Daily:	1.0	Weeks
Weekly:	2.0	Weeks
Monthly:	0.0	Weeks

Lag

Warning Threshold:	1.5	Days
Error Threshold:	2.0	Days

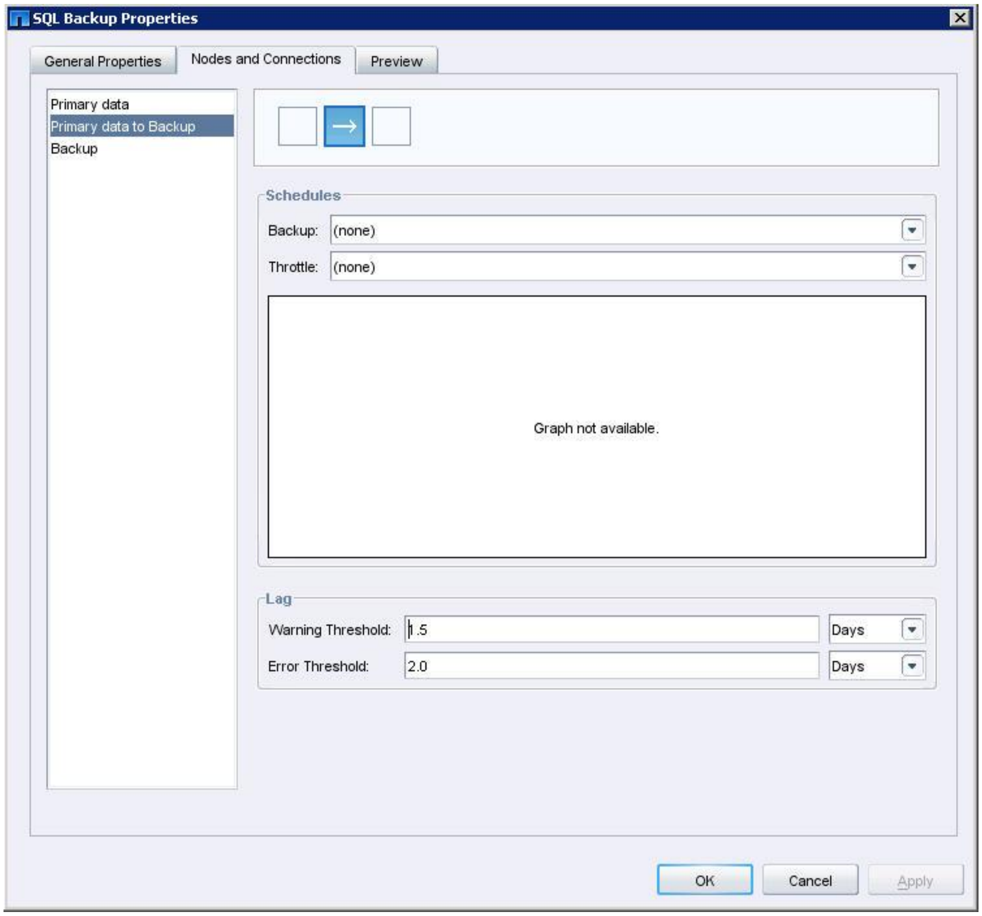
Backup script

Path:

Run As:

Figure 3-2) Dataset properties—primary data.

2. The schedules for both backup and throttle of the primary data should be set to None (Figure 3-3). As part of the backup job created by SnapManager for SQL Server, it is also specified whether or not that Snapshot copy created will be archived with SnapVault. Once SnapManager for SQL Server has finished creating the Snapshot copy, it communicates with Protection Manager to initiate a SnapVault transfer. The warning threshold and error threshold should also be modified to reflect the relevant thresholds. Setting these thresholds too low or too high may not correctly report potential problems in the environment.



Schedules

Backup: (none)

Throttle: (none)

Graph not available.

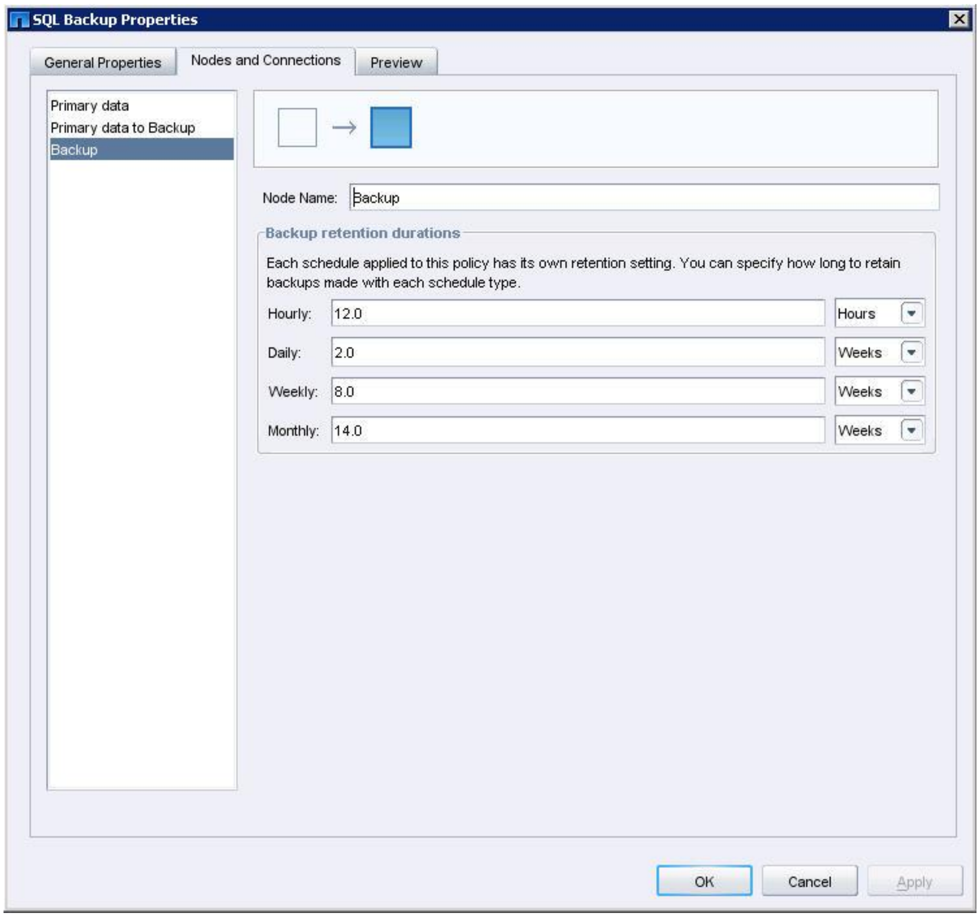
Lag

Warning Threshold: 1.5 Days

Error Threshold: 2.0 Days

Figure 3-3) Primarydata to back up.

3. On the secondary system (designated as Backup in Figure 3-4), the correct retention values need to be specified. These values are dictated by the recovery point objective (RPO) defined by the business in the SLAs. In most cases, the secondary system is where the longer term retention is kept. In this scenario, hourly Snapshot copies are retained for 12 hours; daily Snapshot copies are retained for 2 weeks, weekly for 8 weeks, and monthly for 14 weeks.



SQL Backup Properties

Primary data
Primary data to Backup
Backup

Node Name: Backup

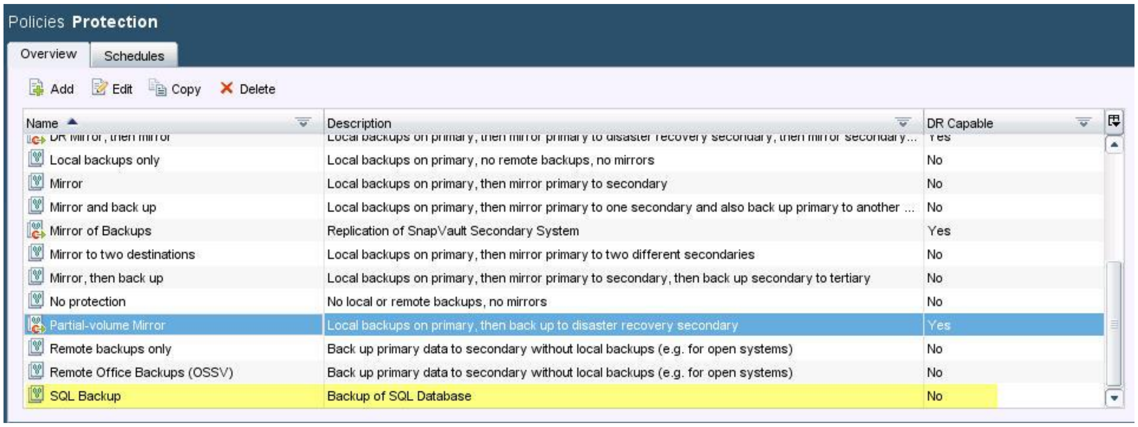
Backup retention durations

Each schedule applied to this policy has its own retention setting. You can specify how long to retain backups made with each schedule type.

Hourly:	12.0	Hours
Daily:	2.0	Weeks
Weekly:	8.0	Weeks
Monthly:	14.0	Weeks

Figure 3-4) Backupretention.

Note: The retention periods defined here are actual time-based retention periods. This is different than the scheduling of SnapVault transfers when using Data ONTAP. Data ONTAP is based on the number of Snapshot copies with a given name.

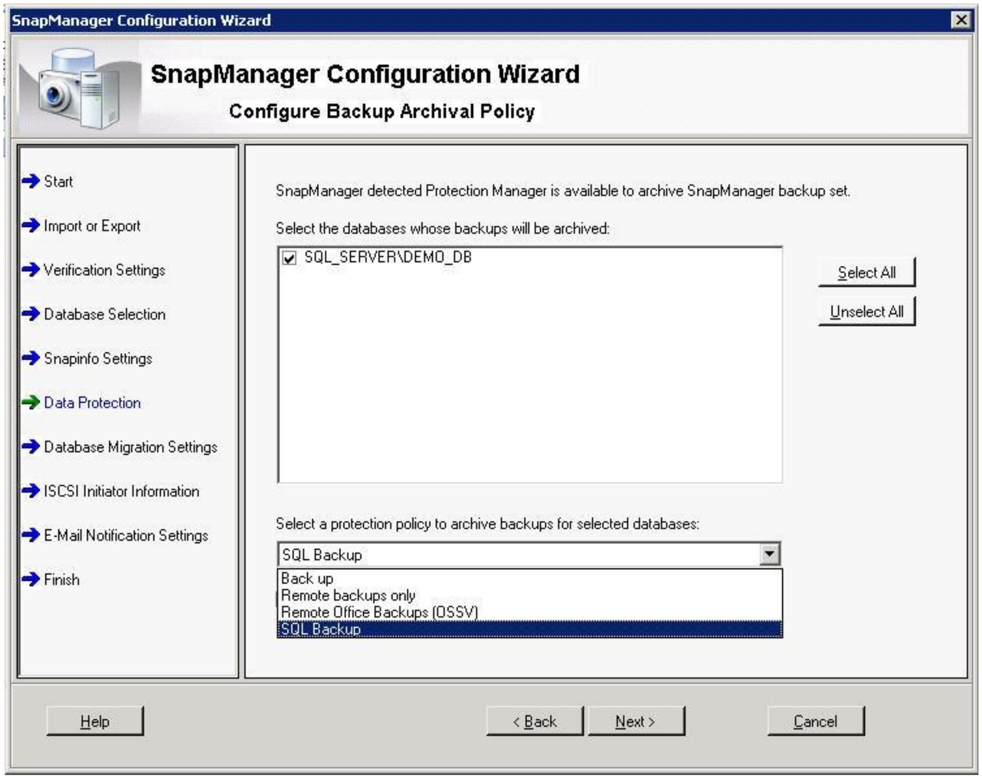


Policies **Protection**

Name	Description	DR Capable
DR Mirror, then mirror	Local backups on primary, then mirror primary to disaster recovery secondary, then mirror secondary...	Yes
Local backups only	Local backups on primary, no remote backups, no mirrors	No
Mirror	Local backups on primary, then mirror primary to secondary	No
Mirror and back up	Local backups on primary, then mirror primary to one secondary and also back up primary to another ...	No
Mirror of Backups	Replication of SnapVault Secondary System	Yes
Mirror to two destinations	Local backups on primary, then mirror primary to two different secondaries	No
Mirror, then back up	Local backups on primary, then mirror primary to secondary, then back up secondary to tertiary	No
No protection	No local or remote backups, no mirrors	No
Partial-volume Mirror	Local backups on primary, then back up to disaster recovery secondary	Yes
Remote backups only	Back up primary data to secondary without local backups (e.g. for open systems)	No
Remote Office Backups (OSSV)	Back up primary data to secondary without local backups (e.g. for open systems)	No
SQL Backup	Backup of SQL Database	No

Figure 3-5) Protection policies.

Once the policy has been created and modified (Figure 3-5), the next step is to configure SnapManager for SQL Server for the backup and replication of the SQL database. In order to replicate a SQL database backup with SnapVault, only one additional step is required during the configuration of the SQL Server through SnapManager for SQL Server. As shown in Figure 26, on the Configure Backup Archival Policy screen, it is required to check the SQL databases that are to be replicated with SnapVault. In this scenario, only one SQL database (DEMO_DB) will be replicated with SnapVault. After the database is selected, the Protection Policy created earlier is selected as the appropriate backup policy for this configuration.


SnapManager Configuration Wizard
SnapManager Configuration Wizard
Configure Backup Archival Policy
Start
Import or Export
Verification Settings
Database Selection
SnapInfo Settings
Data Protection
Database Migration Settings
ISCSI Initiator Information
E-Mail Notification Settings
Finish
Select All
Unselect All
Help
< Back
Next >
Cancel

SnapManager detected Protection Manager is available to archive SnapManager backup set.

Select the databases whose backups will be archived:

☑ SQL_SERVER\DEMO_DB

Select a protection policy to archive backups for selected databases:

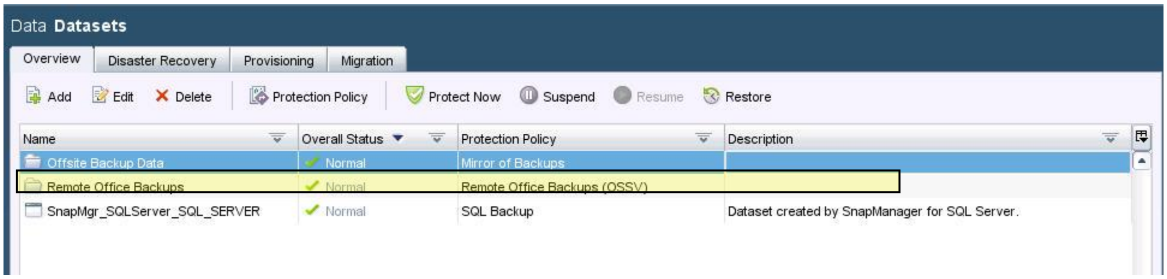
SQL Backup

Back up
Remote backups only
Remote Office Backups (OSSV)
SQL Backup

Figure 3-6) SnapManager SQL data protection.

Note: For more information about configuring SnapManager for SQL Server, see *TR-3768,SnapManager 5.0 for SQL Server Best Practices Guide*.

When the configuration for the backup of the SQL database is complete, a data set automatically appears in Protection Manager (Figure 3-7).



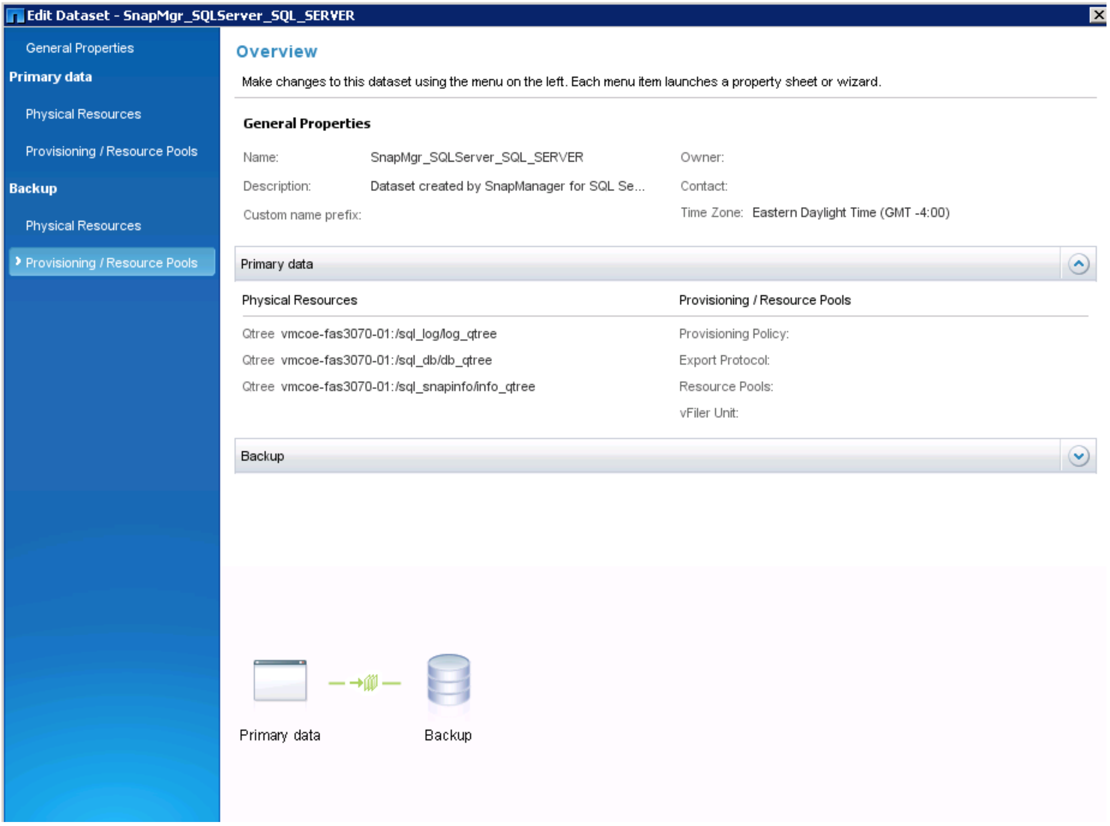
Name	Overall Status	Protection Policy	Description
Offsite Backup Data	✓ Normal	Mirror of Backups	
Remote Office Backups	✓ Normal	Remote Office Backups (OSSV)	
SnapMgr_SQLServer_SQL_SERVER	✓ Normal	SQL Backup	Dataset created by SnapManager for SQL Server.

Figure 3-7) SnapManager for SQL Serverdata set.

The final step in configuring D2D backups for this SQL environment is to assign a resource pool to the dataset. Protection Manager then manages the creation of the destination volumes and performs the first baseline or “level 0” transfer of the SQL database.

To assign a resource pool to a dataset, select the data set and then click on Edit at the top of the window. In the Edit Dataset window that opens, click Provisioning/Resource Pools under Backup (Figure 3-8). Then select the resource pool to associate with this backup. When the resource pool is added to the data set, Protection Manager starts the initial level 0 (baseline) transfers.

Note: This baseline transfer is used only to seed the data to the secondary system; this is not a valid SQL backup because SnapManager for SQL Server did not initiate the backup.



Make changes to this dataset using the menu on the left. Each menu item launches a property sheet or wizard.

General Properties

Name: SnapMgr_SQLServer_SQL_SERVER

Description: Dataset created by SnapManager for SQL Se...

Custom name prefix:

Owner:

Contact:

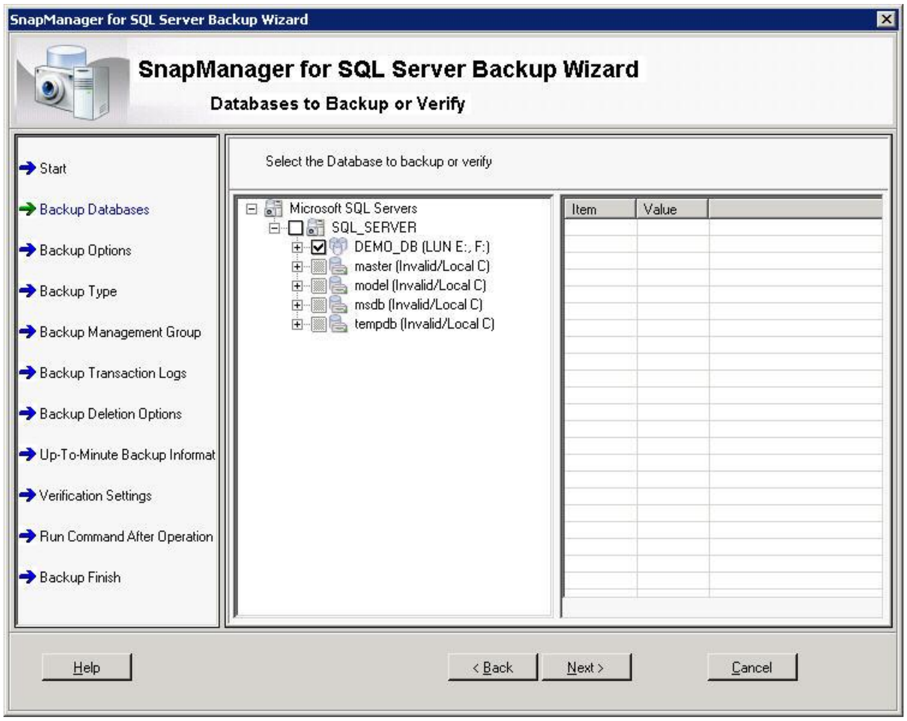
Time Zone: Eastern Daylight Time (GMT -4:00)

Physical Resources	Provisioning / Resource Pools
Qtree vmcoe-fas3070-01:/sql_log/log_qtree	Provisioning Policy:
Qtree vmcoe-fas3070-01:/sql_db/db_qtree	Export Protocol:
Qtree vmcoe-fas3070-01:/sql_snapinfo/info_qtree	Resource Pools:
	vFiler Unit:

Figure 3-8) Dataset properties.

SnapManager for SQL server: creating and scheduling backups

Once theRPOs have been identified for the SQL database, the backup schedules can be created. When configuring SnapManager for SQL Server backups, all backup schedules are initiated from within SQL as a SQL job. SnapManager for SQL Server is used to configure what will be backed up; how long the backup data will be retained locally, and whether or not that backup data will be replicated with SnapVault. The SQL backup jobs can be configured by starting the Backup Wizard from the SnapManager for SQL Server interface. When the Backup Wizard launches, the first step is to select the database to be backed up. This examplebacks up the DEMO_DB database (Figure 3-9). This database spans two LUNs, one for the database and one for the logs.


SnapManager for SQL Server Backup Wizard

SnapManager for SQL Server Backup Wizard

Databases to Backup or Verify

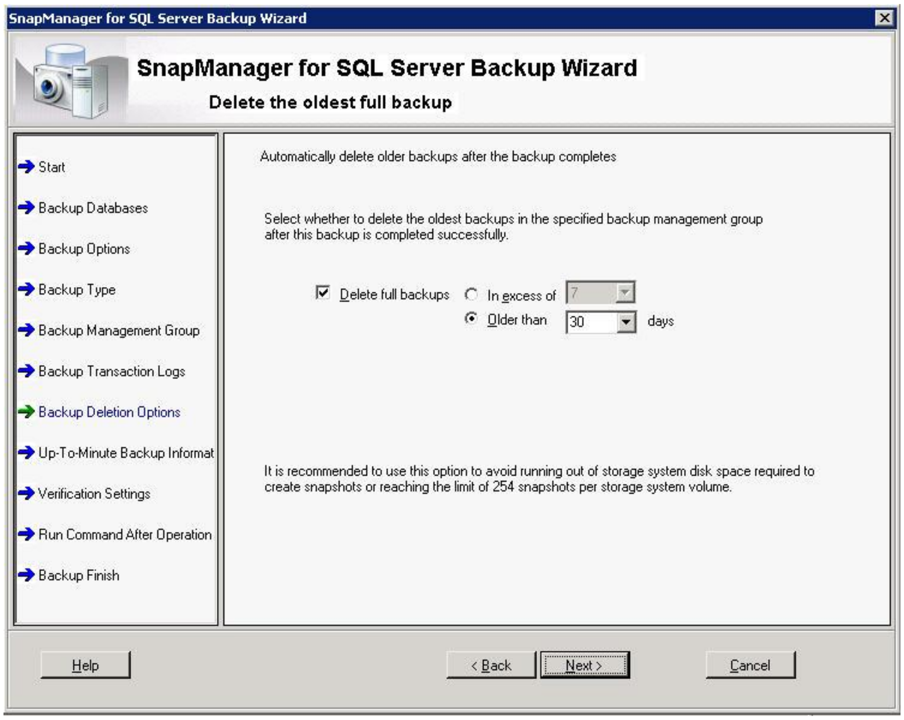
Select the Database to backup or verify

- Microsoft SQL Servers
  - □ SQL_SERVER
    - ☑ DEMO_DB (LUN E:, F:)
    - master (Invalid/Local C)
    - model (Invalid/Local C)
    - msdb (Invalid/Local C)
    - tempdb (Invalid/Local C)

Item	Value

Figure 3-9) Database backup selection.

After the database is selected, various options are available. For details about these options, see *TR-3768:SnapManager for SQL* Server *Best Practices Guide*. This section covers two options. The first option is the ability to delete backups on the local system. SnapManager for SQL Server is responsible for maintaining the appropriate number of local backups in this configuration. In this scenario, only 30 days worth of backups are retained locally. Because this data is being replicated remotely with SnapVault, the data can be stored remotely if a longer retention is required. In addition, because the data is being replicated remotely, it may be acceptable to retain only a minimal number of Snapshot copies locally.


SnapManager for SQL Server Backup Wizard

SnapManager for SQL Server Backup Wizard

Delete the oldest full backup

Automatically delete older backups after the backup completes

Select whether to delete the oldest backups in the specified backup management group after this backup is completed successfully.

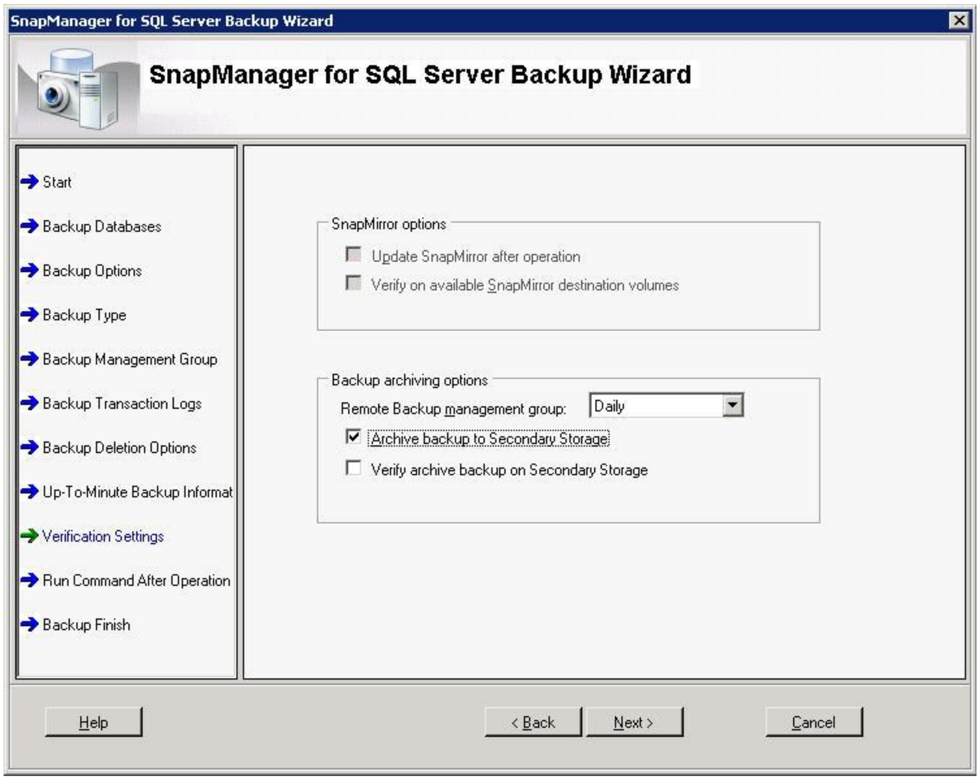
☑ Delete full backups ○ In excess of 7

● Older than 30 days

It is recommended to use this option to avoid running out of storage system disk space required to create snapshots or reaching the limit of 254 snapshots per storage system volume.

Figure 3-10) SnapManager SQL Server local retention.

In this scenario, the SQL Snapshot copies are replicated with SnapVault, so as part of the backup job configuration, it is necessary to select Archive Backup to Secondary Storage. When this box is selected, SnapManager for SQL Server communicates with Protection Manager and replicates the Snapshot copy created on the local storage system. When theoption to archive this backup is selected, the Management Group is also specified. This flags the backup and applies the appropriate retention to the backup. If it is necessary to retain different backup types (daily, hourly, weekly, monthly), then a separate backup job must be scheduled for each.



SnapMirror options

□ Update SnapMirror after operation

□ Verify on available SnapMirror destination volumes

Backup archiving options

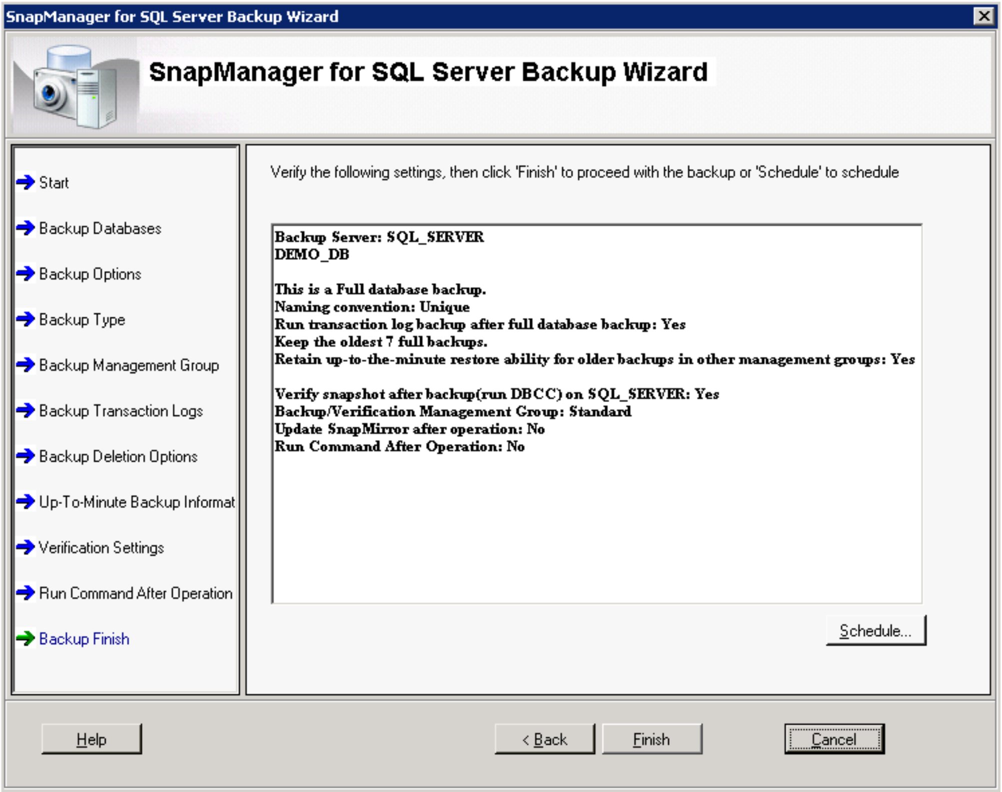
Remote Backup management group: Daily

☑ Archive backup to Secondary Storage

☐ Verify archive backup on Secondary Storage

Figure 3-11) SnapManager for SQL Server archive options.

When all the options have been configured, a summary screen is displayed. Once the summary is validated, two options are available: Click Schedule to schedule the backups to run on a recurring basis; or click Finish to run the backup immediately. If a backup is configured to be replicated by SnapVault, a job is started in Protection Manager and appears as an on-demand backup.



Verify the following settings, then click 'Finish' to proceed with the backup or 'Schedule' to schedule

```
Backup Server: SQL_SERVER
DEMO_DB

This is a Full database backup.
Naming convention: Unique
Run transaction log backup after full database backup: Yes
Keep the oldest 7 full backups.
Retain up-to-the-minute restore ability for older backups in other management groups: Yes

Verify snapshot after backup(run DBCC) on SQL_SERVER: Yes
Backup/Verification Management Group: Standard
Update SnapMirror after operation: No
Run Command After Operation: No
```

Figure 3-12) SnapManager for SQL Server backup summary.

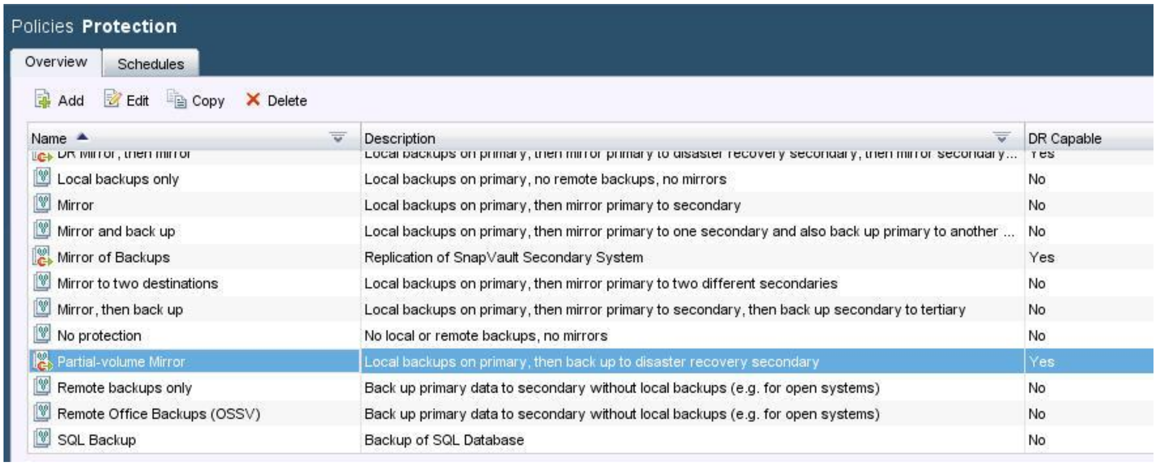
When the schedule has been created, the configuration of the SQL backups is complete. The local backup progress can be monitored from within SnapManager for SQL Server,and the replication of the backups can be monitored through Protection Manager. Another option when using SnapVault and SnapManager for SQL Server is the ability to perform remote verification of the backup data by using a SQL Server and the replicated backup at the centralized data center.

Open Systems SnapVault: configuration

In this portion of the scenario, three Windows servers in the remote office are being protected by Open Systems SnapVault. By using Open Systems SnapVault in the remote office, tape can be eliminated and all backup data is centralized on the same storage system as the SQL backup data.

The first step in this configuration is the installation of the Open Systems SnapVault agent on the remote servers. For more information around the installation and configuration, see the *Open Systems SnapVault Installation and Administration Guide*and*TR-3466, Open Systems SnapVault Best Practices Guide*. After the Open Systems SnapVault agent has been installed on the remote servers and added to the Protection Manager configuration, the Protection Policy must be created. The default Protection Policy, named Remote Backups Only, can be used as the template for the Open Systems SnapVault backup policy.

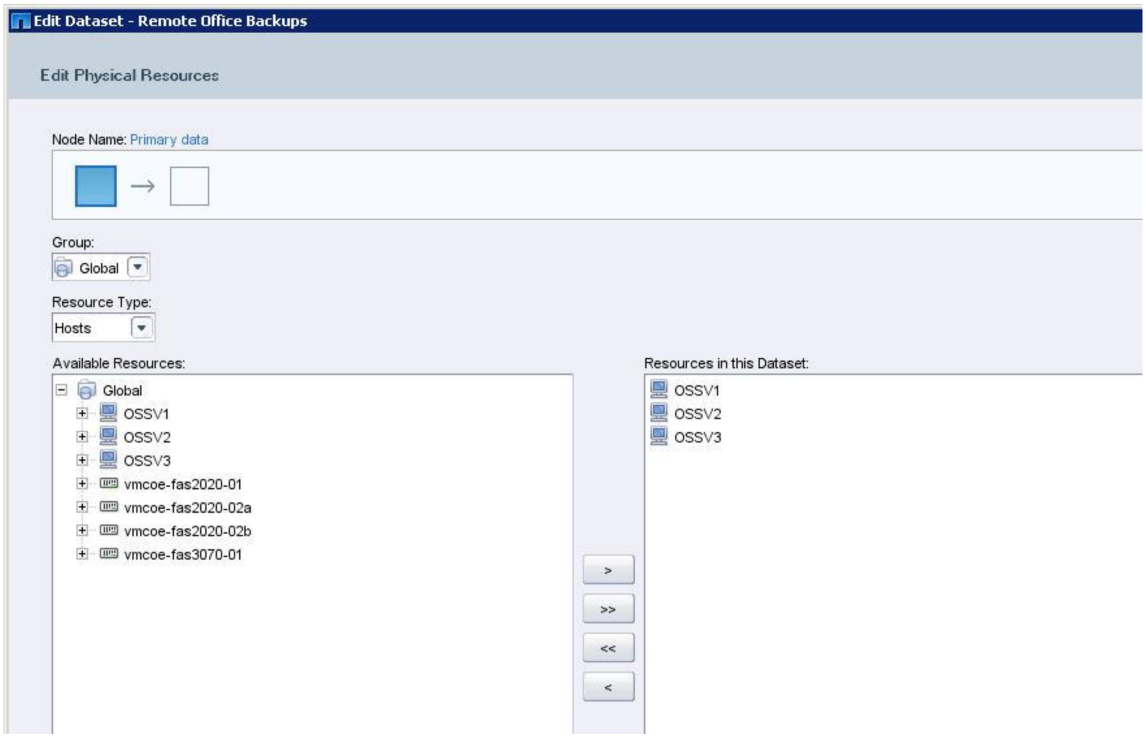
To create a new Protection Policy, select and copy the Remote Backups Only policy. It is important to rename the new Protection Policy and to modify the transfer windows. When modifying the Protection Policy, make sure that a local backup policy is not applied. Because the Open Systems SnapVault clients are not running Data ONTAP, It is not possible to create local Snapshot copies on those systems. Also, there is an optional Throttle policy that can be defined if required; it can be a bandwidth-based throttle or a time-based throttle. If a time-based throttle is applied, backups (either scheduled or adhoc) cannot be executed during this window.



Name	Description	DR Capable
DR Mirror, then mirror	Local backups on primary, then mirror primary to disaster recovery secondary, then mirror secondary...	Yes
Local backups only	Local backups on primary, no remote backups, no mirrors	No
Mirror	Local backups on primary, then mirror primary to secondary	No
Mirror and back up	Local backups on primary, then mirror primary to one secondary and also back up primary to another ...	No
Mirror of Backups	Replication of SnapVault Secondary System	Yes
Mirror to two destinations	Local backups on primary, then mirror primary to two different secondaries	No
Mirror, then back up	Local backups on primary, then mirror primary to secondary, then back up secondary to tertiary	No
No protection	No local or remote backups, no mirrors	No
Partial-volume Mirror	Local backups on primary, then back up to disaster recovery secondary	Yes
Remote backups only	Back up primary data to secondary without local backups (e.g. for open systems)	No
Remote Office Backups (OSSV)	Back up primary data to secondary without local backups (e.g. for open systems)	No
SQL Backup	Backup of SQL Database	No

Figure 3-13) Open Systems SnapVault Protection Policy.

Once the required Protection Policy is created, a data set can be created for the remote office clients. This can be done from theDataset tab in Protection Manager. In this scenario, all remote office clients are grouped into a single dataset. When specifying the primary data in the dataset, the entire client, the entire drive, or individual directories can be selected. If the entire client is selected (Figure 3-14), any new drives or directories added to the individual clients are automatically discovered and protected by Open Systems SnapVault and Protection Manager.


Edit Dataset - Remote Office Backups

Edit Physical Resources

Node Name: Primary data

Group: Global

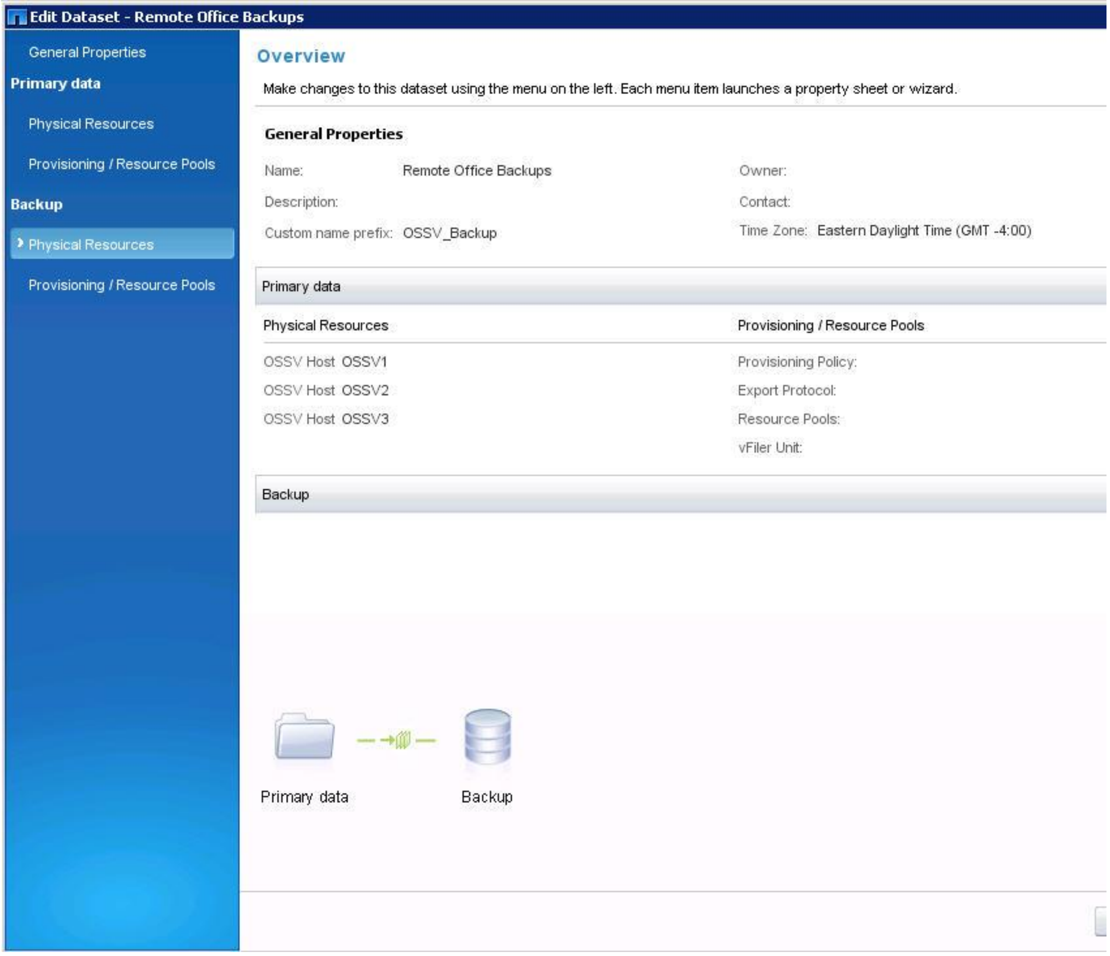
Resource Type: Hosts

Available Resources:
- Global
  - OSSV1
  - OSSV2
  - OSSV3
  - vmcoe-fas2020-01
  - vmcoe-fas2020-02a
  - vmcoe-fas2020-02b
  - vmcoe-fas3070-01

Resources in this Dataset:
- OSSV1
- OSSV2
- OSSV3

Figure 3-14) Open Systems SnapVault primary data.

After the data set has been configured with the appropriate clients and protection policy, the final step is to assign the destination flexible volumes or a resource pool. For ease of monitoring and management, NetApp recommendsthe use of resource pools in Protection Manager. By leveraging this functionality, Protection Manager handles the creation and maintenance of the destination volumes. In order to specify the resource pool, select the appropriate data set and click Edit (Figure 3-15).


Edit Dataset - Remote Office Backups
General Properties
Primary data
Physical Resources
Provisioning / Resource Pools
Backup
Physical Resources
Provisioning / Resource Pools
Primary data
Backup

Overview

Make changes to this dataset using the menu on the left. Each menu item launches a property sheet or wizard.

General Properties

Name: Remote Office Backups

Description:

Custom name prefix: OSSV_Backup

Owner:

Contact:

Time Zone: Eastern Daylight Time (GMT -4:00)

Primary data

Physical Resources

OSSV Host OSSV1

OSSV Host OSSV2

OSSV Host OSSV3

Provisioning / Resource Pools

Provisioning Policy:

Export Protocol:

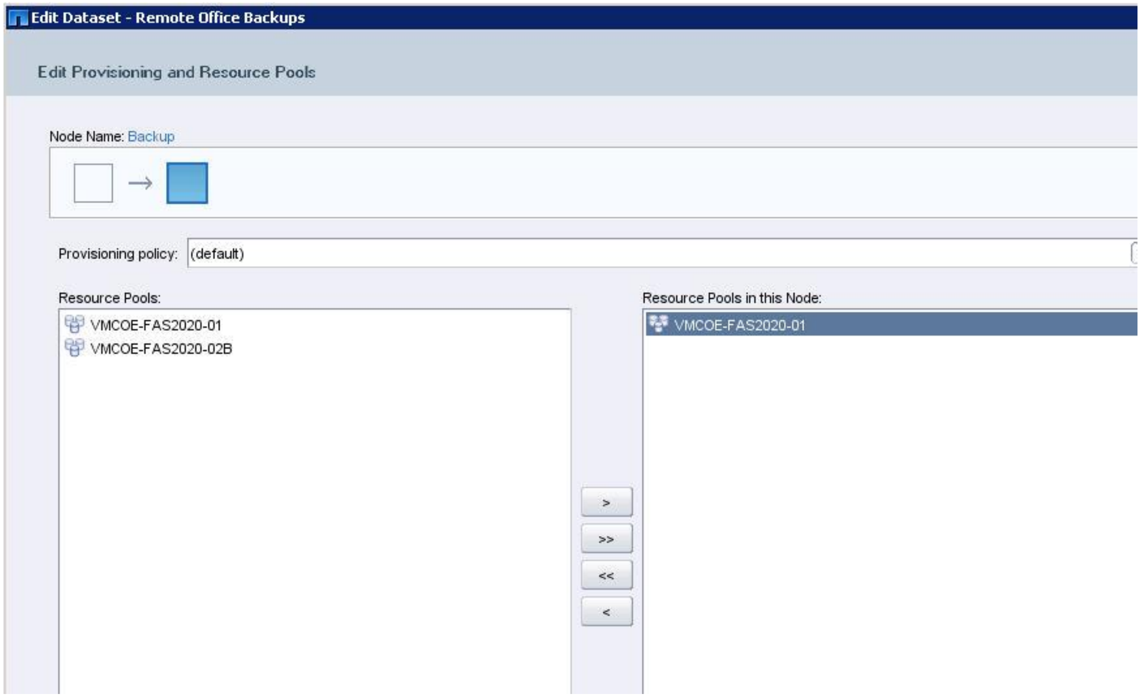
Resource Pools:

vFiler Unit:

Backup

Figure3-15) Open Systems SnapVault: Edit Dataset window.

In the Edit Dataset window, under Backup, click Provisioning/Resource Pools. In the window that opens (Figure 3-16), you can apply the required resource pool to the data set. Protection Manager then automatically creates the flexible volume and performs the initial baseline (level 0) transfer with Open Systems SnapVault.



Edit Dataset - Remote Office Backups

Edit Provisioning and Resource Pools

Node Name: Backup

Provisioning policy: (default)

Resource Pools:
- VMCOE-FAS2020-01
- VMCOE-FAS2020-02B

Resource Pools in this Node:
- VMCOE-FAS2020-01

Figure 3-16) –Open Systems SnapVault: Assign resource pool.

Offsite replication of backup data

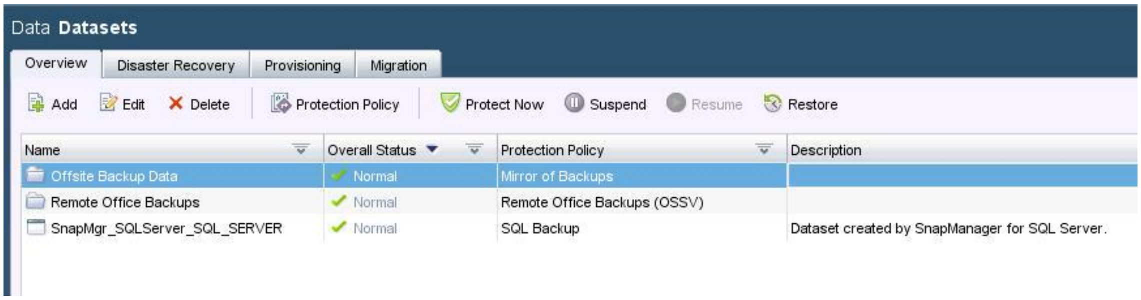
When all the backups have been configured, the final step is to replicate all the backup data to an alternate site. By replicating all the backup data to a remote data center, tape could potentially be completely eliminated as a requirement. The replication of the backup data with volume SnapMirror should satisfy any offsite requirement, assuming that tape (either physical or virtual) isn't required. If tape is required (see the following optional configurations), the tape creation can be shifted from the production storage array to either the backup destination array or the offsite array. When the tape creation is shifted to either of these two systems, the production system is relieved of the tape requirement, so that the production system needs to serve only the production workload.

Protection Manager allows the offsite replication of all the backup data to be configured and monitored in a matter of minutes. The first step is to create a data set of all the backup data. In this example, the data is replicated offsite twice a day, but it can be replicated as required, more frequently or less frequently.

In the data set for the replication of the backup data, it is easiest to select all the backup volumes (both the SQL and remote office data) so that only a single data set needs to be monitored and managed. If new data is added to the backup system with an offsite requirement, it can easily be added to the data set, and Protection Manager will handle the replication of that data.

Another advantage of replicating all the backup data to an alternate location is the ability to continue the backup replication if something happens to the secondary storage system. To achieve this, the SnapVault and Open Systems SnapVault relationships must be modified to point directly to the tertiary systems, but another baseline (level 0) copy is not required.

After the [data set for the offsite replication is created, there should be three data sets in Protection Manager (Figure 3-17). For larger environments there may be more datasets, but the overall strategy and configuration will be very similar.



Data **Datasets**

Name	Overall Status	Protection Policy	Description
Offsite Backup Data	✓ Normal	Mirror of Backups	
Remote Office Backups	✓ Normal	Remote Office Backups (OSSV)	
SnapMgr_SQLServer_SQL_SERVER	✓ Normal	SQL Backup	Dataset created by SnapManager for SQL Server.

Figure 3-17) Dataset summary.

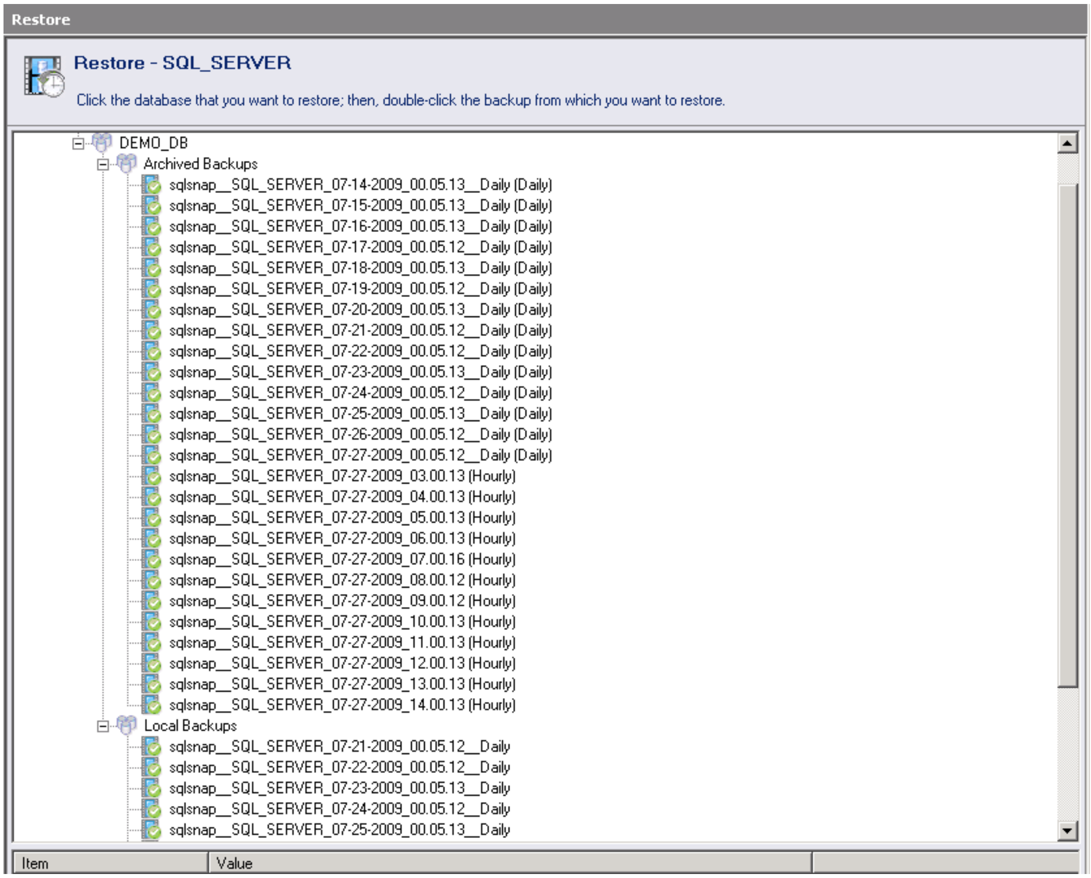
The configuration of the environment for this scenario is now complete. Protection Manager can now be configured to monitor the environment and send alerts based on user-defined variables and settings. NetApp recommends that the progress and success of the data replication transfers be monitored through Protection Manager on a regular basis.

The next section of this scenario discusses some common recovery methods and how to complete the recovery of the data.

RECOVERY METHODS

Recovering SQL local or remote copy

Backup of the SQL database is important, but recovery is where the true test of the solution lies. When recovering a SQL database, recovery is always initiated from within SnapManager for SQLServer. When recovering from a backup of the SQL Server, it is possible to choose local or remote (archived) backup. The overall restore process is the same, except for data being moved from the secondary to the primary system if an Archived Backup is used for the restore. To view all possible recovery points available for restore, select Restore under the SQL Server instance in SnapManager for SQL Server.



Restore

Restore - SQL_SERVER

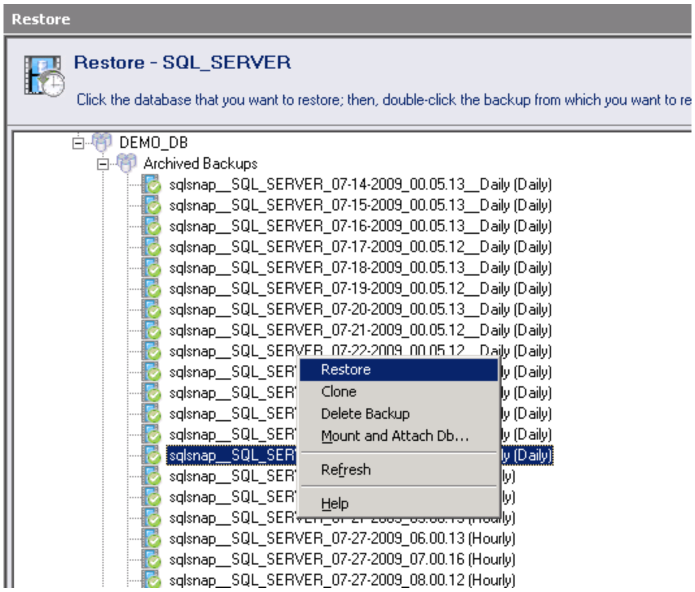
Click the database that you want to restore; then, double-click the backup from which you want to restore.

```
DEMO_DB
  Archived Backups
    sqlsnap__SQL_SERVER_07-14-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-15-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-16-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-17-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-18-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-19-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-20-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-21-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-22-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-23-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-24-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-25-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-26-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-27-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-27-2009_03.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_04.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_05.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_06.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_07.00.16 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_08.00.12 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_09.00.12 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_10.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_11.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_12.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_13.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_14.00.13 (Hourly)
  Local Backups
    sqlsnap__SQL_SERVER_07-21-2009_00.05.12__Daily
    sqlsnap__SQL_SERVER_07-22-2009_00.05.12__Daily
    sqlsnap__SQL_SERVER_07-23-2009_00.05.13__Daily
    sqlsnap__SQL_SERVER_07-24-2009_00.05.12__Daily
    sqlsnap__SQL_SERVER_07-25-2009_00.05.13__Daily
```

Item	Value

Figure 3-18) SnapManager for SQL Server Restore window.

The first step in performing a restore from SnapManager for SQL Server is to determine the backup that needs to be restored. Once the backup is determined, either openthe Restore Wizard or double-click the backup instance to be restored and then right-click the instance and select Restore. If the data is to be restored from an alternate location, such as tape, then the Restore Wizard should be used.

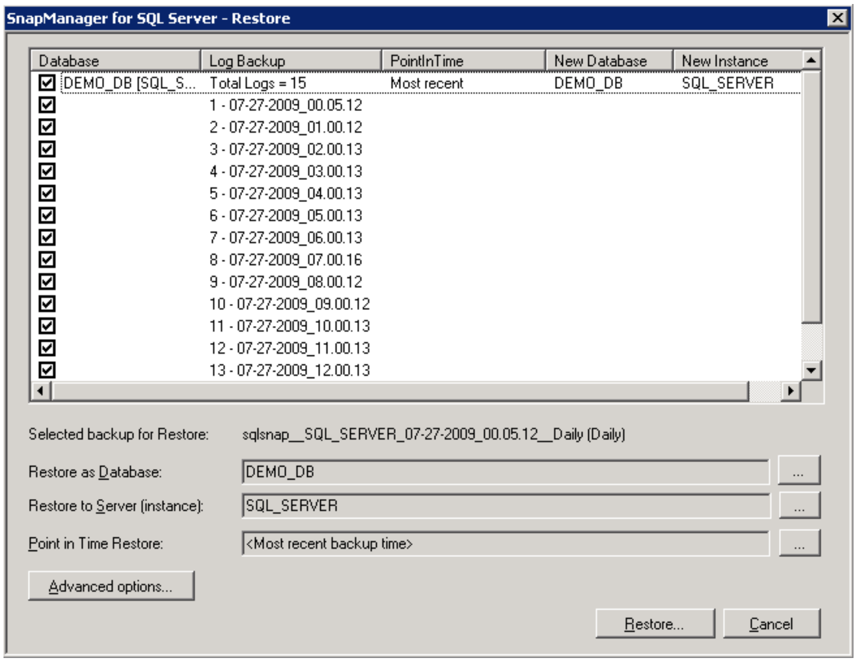


Click the database that you want to restore; then, double-click the backup from which you want to re

```
DEMO_DB
  Archived Backups
    sqlsnap__SQL_SERVER_07-14-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-15-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-16-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-17-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-18-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-19-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-20-2009_00.05.13__Daily (Daily)
    sqlsnap__SQL_SERVER_07-21-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SERVER_07-22-2009_00.05.12__Daily (Daily)
    sqlsnap__SQL_SER                                  ly (Daily)
    sqlsnap__SQL_SER                                  ly (Daily)
    sqlsnap__SQL_SER                                  ly (Daily)
    sqlsnap__SQL_SER                                  ly (Daily)
    sqlsnap__SQL_SER                                  ly (Daily)
    sqlsnap__SQL_SER                                  ly)
    sqlsnap__SQL_SER                                  ly)
    sqlsnap__SQL_SER                                  ly)
    sqlsnap__SQL_SERVER_07-27-2009_06.00.13 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_07.00.16 (Hourly)
    sqlsnap__SQL_SERVER_07-27-2009_08.00.12 (Hourly)
```

Figure 3-19) Selecting a recovery point.

When the backup to restore is selected, the Restore window opens, in whicha specific point in time can be selected for the restore. By default, SnapManager for SQL Server restores to the most recent backup time. After selecting a point in time, click Restore. A window opens with three options for the restore settings. By default, SnapManager for SQL Server leaves the databases operational but unavailable for restoring additional transaction logs.



Database	Log Backup	PointInTime	New Database	New Instance
☑ DEMO_DB [SQL_S...	Total Logs = 15	Most recent	DEMO_DB	SQL_SERVER
☑	1 - 07-27-2009_00.05.12			
☑	2 - 07-27-2009_01.00.12			
☑	3 - 07-27-2009_02.00.13			
☑	4 - 07-27-2009_03.00.13			
☑	5 - 07-27-2009_04.00.13			
☑	6 - 07-27-2009_05.00.13			
☑	7 - 07-27-2009_06.00.13			
☑	8 - 07-27-2009_07.00.16			
☑	9 - 07-27-2009_08.00.12			
☑	10 - 07-27-2009_09.00.12			
☑	11 - 07-27-2009_10.00.13			
☑	12 - 07-27-2009_11.00.13			
☑	13 - 07-27-2009_12.00.13			

Selected backup for Restore: sqlsnap__SQL_SERVER_07-27-2009_00.05.12__Daily (Daily)

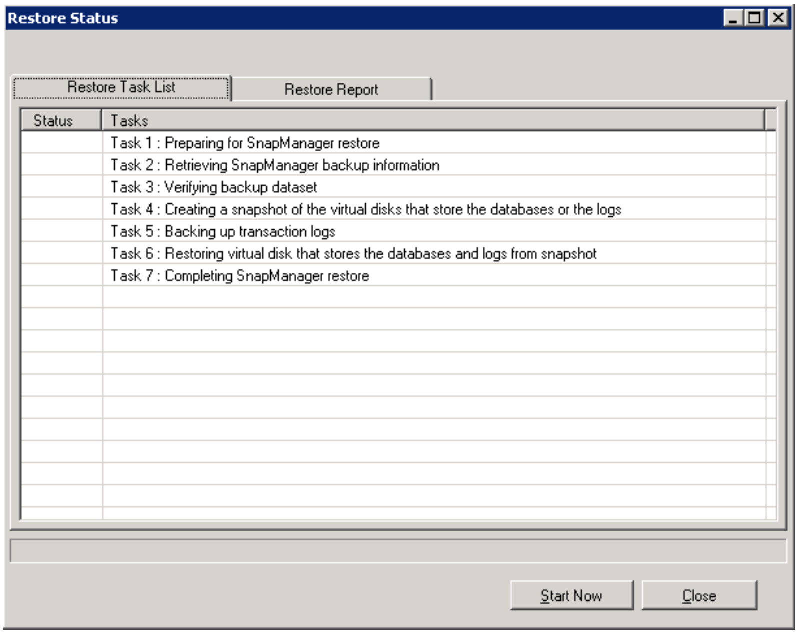
Restore as Database: DEMO_DB

Restore to Server (instance): SQL_SERVER

Point in Time Restore: <Most recent backup time>

Figure 3-20) Restore window.

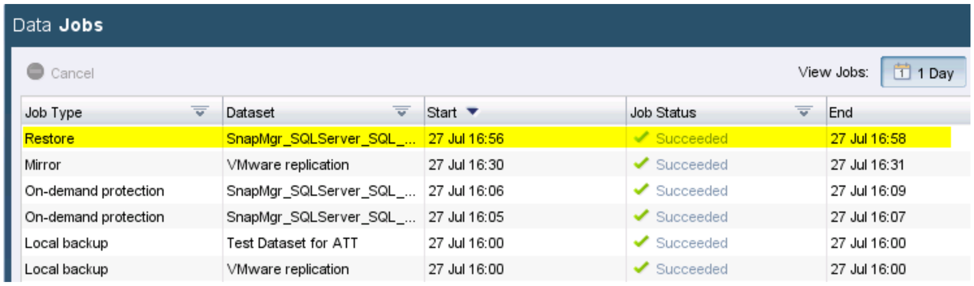
After the restore options have been set, theRestore Status window opens (Figure 3-20). Click Start Now to begin the restore. During the restore process, the current status of the restore is displayed.



Status	Tasks
	Task 1 : Preparing for SnapManager restore
	Task 2 : Retrieving SnapManager backup information
	Task 3 : Verifying backup dataset
	Task 4 : Creating a snapshot of the virtual disks that store the databases or the logs
	Task 5 : Backing up transaction logs
	Task 6 : Restoring virtual disk that stores the databases and logs from snapshot
	Task 7 : Completing SnapManager restore

Figure 3-21) Restore Status window.

When restoring from an archived backup, SnapManager for SQL Server contacts Protection Manager and starts a restore job. If any portion of the data still resides on disk, SnapVault performs an incremental restore, which transfer only the blocks required to recover to the specified point in time. The progress of the SnapVault transfer can be monitored from Protection Manager or by running a `snapvault status` command on the NetApp storage system.



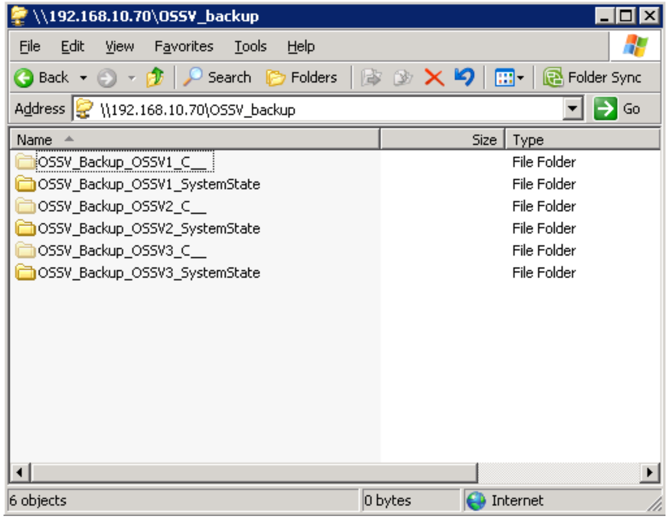
Job Type	Dataset	Start	Job Status	End
Restore	SnapMgr_SQLServer_SQL_...	27 Jul 16:56	✓ Succeeded	27 Jul 16:58
Mirror	VMware replication	27 Jul 16:30	✓ Succeeded	27 Jul 16:31
On-demand protection	SnapMgr_SQLServer_SQL_...	27 Jul 16:06	✓ Succeeded	27 Jul 16:09
On-demand protection	SnapMgr_SQLServer_SQL_...	27 Jul 16:05	✓ Succeeded	27 Jul 16:07
Local backup	Test Dataset for ATT	27 Jul 16:00	✓ Succeeded	27 Jul 16:00
Local backup	VMware replication	27 Jul 16:00	✓ Succeeded	27 Jul 16:00

Figure 3-22) Protection Manager restore.

When the restore has completed successfully, SnapManager for SQL Server mounts and starts the database instance.

Recovering data for remote clients

The recovery of data for remote clients is simplified by using Open Systems SnapVault as the backup solution. Because the data is stored in an open format, the backup data on the secondary or tertiary storage system can be mounted via a CIFS/NFS connection. The most recent backup data is displayed in the root of the volume. If other versions of the data are required, the Snapshot directory can be used (~snapshot on Windows or .snapshot on UNIX or Linux). When the correct version of the file it located, it can be copied back to the restore location.



\\192.168.10.70\OSSV_backup

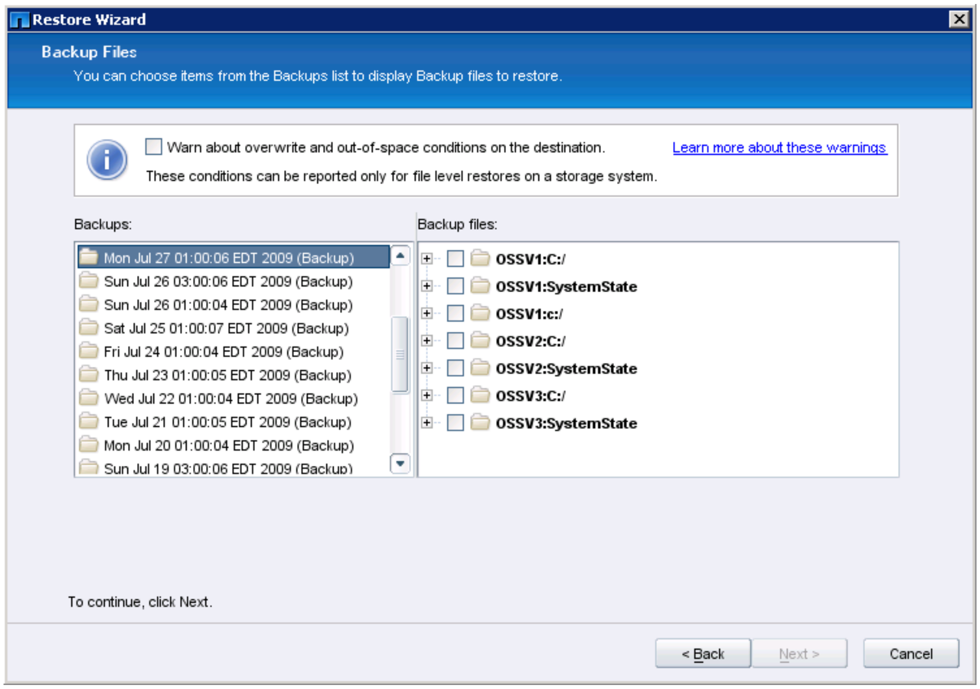
Address: \\192.168.10.70\OSSV_backup

Name	Size	Type
OSSV_Backup_OSSV1_C__		File Folder
OSSV_Backup_OSSV1_SystemState		File Folder
OSSV_Backup_OSSV2_C__		File Folder
OSSV_Backup_OSSV2_SystemState		File Folder
OSSV_Backup_OSSV3_C__		File Folder
OSSV_Backup_OSSV3_SystemState		File Folder

6 objects | 0 bytes | Internet

Figure 3-23) CIFS mount for restore.

Another option for restore is to use the Protection Manager interface. Protection Manager enables simple restore of entire file systems, directories, or even just a few files. To restore the data from Protection Manager, select a dataset and then click Restore at the top of the window (Figure 3-24). When the data to be restored is selected, an optionis available to restore the data in place (overwriting any existing files) or to restore to an alternate location. For an alternate location restore, the data can be restored either to an alternate location on the original backup system or to an alternate system. When the restore has been initiated, a job is created in Protection Manager and can be monitored with the Protection Manager software.



Restore Wizard

Backup Files

You can choose items from the Backups list to display Backup files to restore.

□ Warn about overwrite and out-of-space conditions on the destination. Learn more about these warnings

These conditions can be reported only for file level restores on a storage system.

Backups:

- Mon Jul 27 01:00:06 EDT 2009 (Backup)
- Sun Jul 26 03:00:06 EDT 2009 (Backup)
- Sun Jul 26 01:00:04 EDT 2009 (Backup)
- Sat Jul 25 01:00:07 EDT 2009 (Backup)
- Fri Jul 24 01:00:04 EDT 2009 (Backup)
- Thu Jul 23 01:00:05 EDT 2009 (Backup)
- Wed Jul 22 01:00:04 EDT 2009 (Backup)
- Tue Jul 21 01:00:05 EDT 2009 (Backup)
- Mon Jul 20 01:00:04 EDT 2009 (Backup)
- Sun Jul 19 03:00:06 EDT 2009 (Backup)

Backup files:

- □ OSSV1:C:/
- □ OSSV1:SystemState
- □ OSSV1:c:/
- □ OSSV2:C:/
- □ OSSV2:SystemState
- □ OSSV3:C:/
- □ OSSV3:SystemState

To continue, click Next.

Figure 3-24) Protection Manager restore.

The final method of recovery for remote office systems is to use the CLI on the Open Systems SnapVault clients. The `snapvault restore` command can be used on any Open Systems SnapVault client to perform filesystem, file, or directory level recoveries. Because all Open Systems SnapVault transfers are "pull" transfers, the `snapvault restore` command must be issued from the client to whichthe data will be restored.

OPTIONAL CONFIGURATIONS

Backing up data to tape—virtual or physical

The final consideration in this scenario is whether or not it is acceptable to go 100% tapeless in the environment. Not all environments are able to eliminate tape completely,for reasons such as comfort level and legal requirements. If tape is a requirement in the environment, the tapes should be created from either the secondary or the tertiary storage system. When either of those systems is used to create the tape backups, the production workload (hosts or storage system) is not affected and can continue to serve up the production data without the impact of creating the tapes.

To create the tapes, the tape library unit can be attached directly to the NetApp storage system and the backups can be controlled by the existing tape backup software provider. NDMP can be used as the data transfer protocol to create the tapes from the specified Snapshot copy.

Note:NDMP only transfers the data for a single Snapshot copy (not all Snapshot copies) for a given volume. Also, an NDMP transfermust be controlled through the existing tape backup software; Protection Manager does not support the creation of tapes.

For more information about configuring a tape backup solution with NetApp Storage systems, see the *Data Protection Tape Backup and Recovery Guide* for the version of Data ONTAP running on the NetApp storage system.

3.2 SCENARIO 2: ASYNCHRONOUS MIRRORING AND FLEX CLONE COPIES FOR DEVELOPMENT AND TESTING

Products: SnapMirror Async (Qtree), FlexClone®, Snapshot Copies

OVERVIEW

This scenario describes a disaster recovery setup, meaning that data from the primary production site is replicated to a DR site. This site can be immediately available in the event of a disaster so that users can be redirected to the DR data. During normal production operations, when the primary site is active, the DR site and therefore the DR infrastructure are idle and unused. In traditional deployments, two key challenges arise when customers want to use these idle DR resources. The first is that DR replication must be stopped in order to fully test the DR copies or to use them for database development, testing, and tagging.The second challenge is that each test copy at the DR site takes up space, thus increasing the footprint.

However, with SnapMirror and FlexClone, users can actively use the DR resources for nonproduction purposes such as DR testing, database development, testing, and staging while the primary production site is active and without interrupting replication of critical data to the DR site. Also, the copies at the DR site that are used for testing do not consume any additional space. These space-efficient copies start consuming space only when data is changed or when new data is created. Therefore customers can instantly create many space-efficient copies without increasing the footprint.

The ability to simultaneously use the standby DR resources while making sure that the critical data is replicated is very appealing to many customers. Another important value proposition is the fact that the copies at the DR sites take only seconds to create, with no additional storage.

BUSINESS PROBLEMS AND PROTECTION

Suppose that a customer wants to implement a disaster recovery site but is having a hard time justifying the expense. Financial planners say that the risk is not worth the cost of a dedicated DR location. The solution proposed and implemented allows replication for DR purposes. To further leverage the DR location, FlexClone copies are proposed in order to make nondisruptive clones for test, development, reporting, and training purposes.

RECOVERY METHODS

With asynchronous replication, upon the loss of the production site, Oracle applications can now be recovered in minutes. In addition, because multiple Oracle backups are being replicated, recovery to a specific backup point in time is possible.

IMPLEMENTATION

The following table outlines the process of creation and verification of database clones for development and testing.

Table 3-1) Implementation steps.

Step	Description
1	Install and configure the NetApp storage.including provisioning of storage and installation of licenses.
2	Install and configure the database hosts.
3	Install and configure Oracle, including hot backups.
5	Configure SnapMirror and verify proper Snapshot copies from hot backup and operation of SnapMirror.
8	Wait until next the hot backup and then perform a SnapMirror update from the latest hotbackup.
9	Create a FlexClone copy from the SnapMirror destination and bring up the alternate server on the clone.

INSTALL AND CONFIGURE NETAPP STORAGE

Two FAS3070 storage systems were used for storage of an Oracle11*g*™ database log and control files (Figure 45).

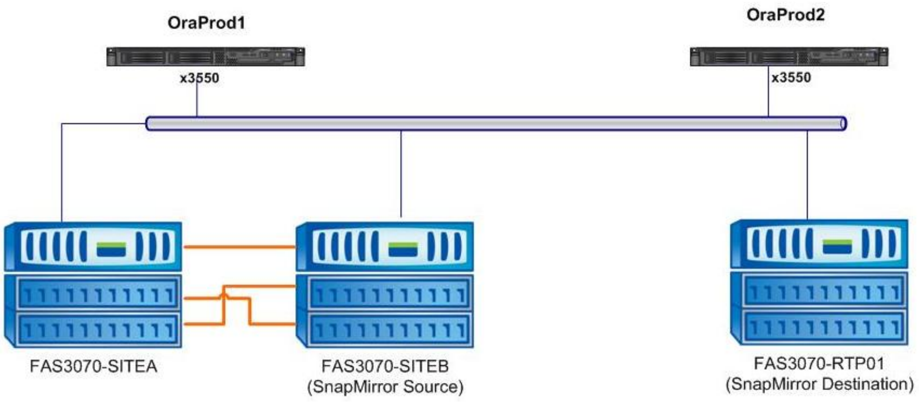


Figure 3-25) Test environment.

Although the primary storage system for Oracle was FAS3070-SITEB, part of a NetApp MetroCluster, MetroCluster functionality was not part of this scenario. FAS3070-RTP01 was used as the replication destination. Table 3-2 lists the specifications of the NetApp FAS systems used.

Table 3-2) NetApp FAS system specifications.

	FAS3070-SITEB	FAS3070-RTP01
Model	FAS3070	FAS3070
Operating System	Data ONTAP 7.2.6.1	Data ONTAP 7.3.1
Licenses	NFS, SnapMirror	NFS, SnapMirror,FlexClone

SOURCE NETAPP STORAGE SYSTEM (FAS3070-SITEB)

Disk aggregates and volumes were created as shown in Figure 46. Two volumes, U03 and U04, were created to house the NFS storage for the Oracle environment.

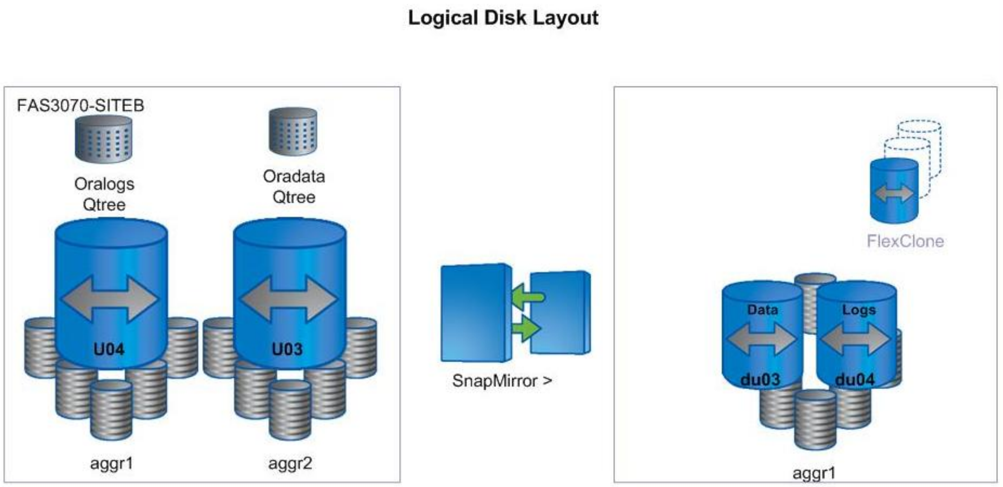


Figure 3-26) Disk layout.

One qtree was created in each volume for the log and control files (Oralogs) and another for the data files (Oradata) Each qtree was added to the NFS exports file and exported for mounting by the Oracle server.

CONFIGURE DESTINATION STORAGE

The disaster recovery FAS storage controller, FAS3070-RTP01, contains the aggregate (aggr1) that will host the SnapMirror replication destination.

There are two destination volumes in this scenario. Volume du03 contains the replicated control and log files. Volume du04 contains the replicated data files.

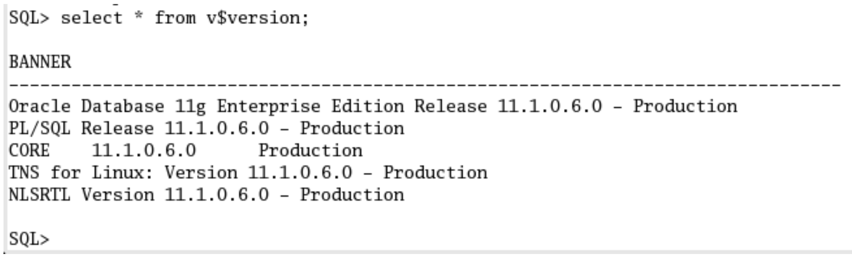
INSTALL AND CONFIGURE HOST SERVERS

For this test scenario, two IBM 3550 servers with 4GB of memory (Oraprod1 and Oraprod2) were used. Each used local direct-attached storage for the operating system and the Oracle binaries. Oraprod1 is the main production database server. It is running Red Hat Advanced server 4.0.

Oraprod2 is the server used to run the clone off of the database replication destination. It is also running Red Hat Advanced server 4.0.

INSTALL AND CONFIGURE ORACLE

Oracle Enterprise Edition was installed on both servers, Oraprod1 and Oraprod2, as shown in Figure 47.

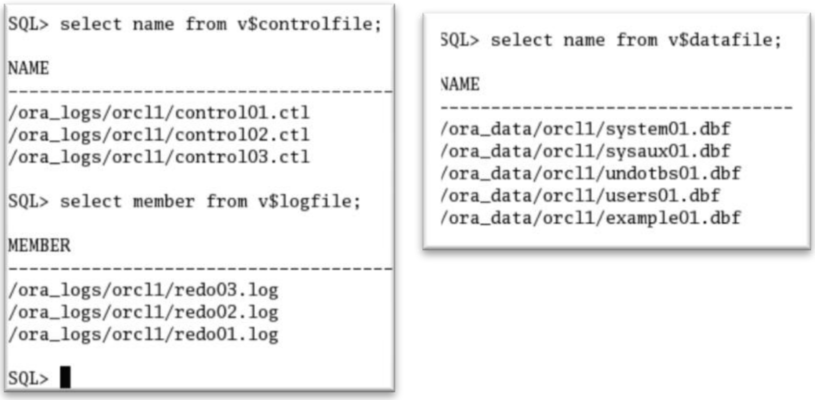


```
SQL> select * from v$version;

BANNER
--------------------------------------------------------------------------------
Oracle Database 11g Enterprise Edition Release 11.1.0.6.0 - Production
PL/SQL Release 11.1.0.6.0 - Production
CORE    11.1.0.6.0      Production
TNS for Linux: Version 11.1.0.6.0 - Production
NLSRTL Version 11.1.0.6.0 - Production

SQL>
```

Figure 3-27) Oracle version information (Oraprod1and Oraprod2).

ORAPROD1

The Oracle data files were placed on the NFS mounted volume called`/Oradata`. The log and control files were located on `/Oralogs`, also an NFS mounted volume (Figure 3-28).



```
SQL> select name from v$controlfile;

NAME
--------------------------------------
/ora_logs/orcl1/control01.ctl
/ora_logs/orcl1/control02.ctl
/ora_logs/orcl1/control03.ctl

SQL> select member from v$logfile;

MEMBER
--------------------------------------
/ora_logs/orcl1/redo03.log
/ora_logs/orcl1/redo02.log
/ora_logs/orcl1/redo01.log

SQL> █
```

```
SQL> select name from v$datafile;

NAME
-----------------------------------
/ora_data/orcl1/system01.dbf
/ora_data/orcl1/sysaux01.dbf
/ora_data/orcl1/undotbs01.dbf
/ora_data/orcl1/users01.dbf
/ora_data/orcl1/example01.dbf
```

Figure 3-28) Database file locations.

INSTALL AND CONFIGURE DATABASE HOT BACKUP SCRIPTS

Scripts were created to perform an Oracle hot backup on a scheduled basis.

CONFIGURE SNAPMIRROR ON SOURCE AND DESTINATION

As previously mentioned, NetApp SnapMirror was licensed on both the source and destination FAS controllers.

VERIFY SUCCESSFUL BACKUP AND REPLICATION OPERATION

Make sure that the hot backup script is running correctly on the production host, Oraprod1. Do this by looking at the Snapshot copies on the production FAS controller. You will also notice the Snapshot copy made (and locked) by SnapMirror.

On the destination FAS controller, check the SnapMirror status of those destination volumes to verify correct operation

PERFORM SNAPMIRROR UPDATE

In this scenario, a manual SnapMirror update was performed. Normally the updates would run on a regularly scheduled basis.

CREATE A FLEXCLONE FROM THE SNAPMIRROR DESTINATION AND BRING IT INTO OPERATION

The following is a description of the clone command that was used to clone the volume containing the Oracle database.

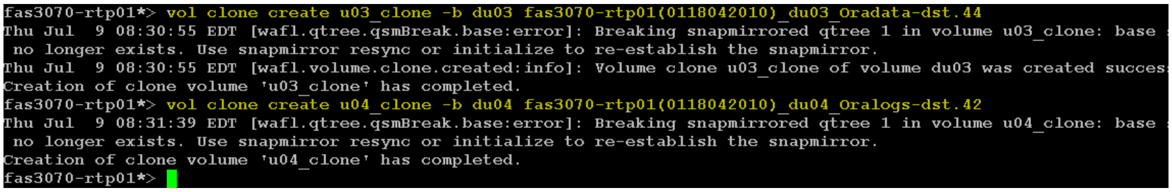
vol clone create *clone_name* [-s {volume|file|none}] -b *parent_name* [*parent_snap*]

`clone_name` is the name of the FlexClone volume that you want to create.

`-s {volume | file | none}` specifies the space guarantee setting for the new FlexClone volume. If no value is specified, the FlexClone volume is given the same space guarantee setting as its parent.

`parent_name` is the name of the FlexVol volume that you intend to clone.

`parent_snap` is the name of the base Snapshot copy of the parent FlexVol volume. If no name is specified, Data ONTAP creates a base Snapshot copy with the name `clone_cl_name_prefix.id`, where `cl_name_prefix` is up to 16 characters of the name of the new FlexClone volume and `id` is a unique digit identifier (for example 1, 2, and so on).



```
fas3070-rtp01*> vol clone create u03_clone -b du03 fas3070-rtp01(0118042010)_du03_Oradata-dst.44
Thu Jul  9 08:30:55 EDT [wafl.qtree.qsmBreak.base:error]: Breaking snapmirrored qtree 1 in volume u03_clone: base
 no longer exists. Use snapmirror resync or initialize to re-establish the snapmirror.
Thu Jul  9 08:30:55 EDT [wafl.volume.clone.created:info]: Volume clone u03_clone of volume du03 was created succes
Creation of clone volume 'u03_clone' has completed.
fas3070-rtp01*> vol clone create u04_clone -b du04 fas3070-rtp01(0118042010)_du04_Oralogs-dst.42
Thu Jul  9 08:31:39 EDT [wafl.qtree.qsmBreak.base:error]: Breaking snapmirrored qtree 1 in volume u04_clone: base
 no longer exists. Use snapmirror resync or initialize to re-establish the snapmirror.
Creation of clone volume 'u04_clone' has completed.
fas3070-rtp01*>
```

Figure 3-29) Creation of clones of the Oracle database volumes.

The message in Figure 3-29 containsa reference to the breaking of a SnapMirror relationship. This is for the clone, not the original volume or qtree. That replication relationship for the primary volume continues uninterrupted.

On the other Oracle server, Oraprod2, check to make sure that the new clone volumes are mounted. If the clone is made of a volume containing an NFS exported qtree, it is automatically added to the exports file.

The database started from the clone volumes can now be used for development, testing, and staging. It is also possible to clone a FlexClone volume. When the testing is complete, the FlexClone volume can either be destroyed or split. If the FlexClone volume is split, the volume will consume space. Therefore, make sure that you have sufficient space in the aggregate before splitting a clone.

SUMMARY AND BEST PRACTICES

Unlike volume SnapMirror, qtree SnapMirror does not maintain the same number of Snapshot copies on source and destination systems. Other than the SnapMirror baseline Snapshot copy, different Snapshot copies can be maintained at primary and destination systems. The advantage of qtree SnapMirror is that a FlexClone volume can live for a long time on the SnapMirror destination system without space implications on the source system. This is not the case with volume SnapMirror.

SnapMirror updates require a common SnapMirror Snapshot copy. Therefore, do not delete SnapMirror Snapshot copies on either the source or the destination system.

FlexClone volumes are backed up the Snapshot copy from which they are created. The backing Snapshot copy is hard-locked (with a Busy tag) and therefore cannot be deleted. When the FlexClone volume is destroyed, the lock is removed as well.

If the FlexClone volume is split, it becomes a normal flexible volume and therefore requires full space allocation, depending on the space guarantees. When the FlexClone volume is split, all existing Snapshot copies of the volume are deleted. For more information, refer to the Storage Management Guide on NOW™ (NetApp on the Web).

1 VERSION HISTORY

Version 1.1	January 2011 Minor updates
Version 1.0	May 2009 Original document

www.netapp.com