Choosing Data Protection Technologies

White Paper

Paul Mayer
Product Manager

August 26, 2002

Table of Contents

Abstract

Decision making about protecting data

In today's information driven organizations, the costs to generate, protect, and manage data are staggering. With the widespread increased reliance upon data, the costs of interrupted access to data or data losses could financially compromise the organization, and in some cases, leave it no room to recover. Given that, decisions pertaining to storing and protecting the organization's data assets can no longer be the exclusive responsibility of IT experts, but must involve discussions with technologically informed CIOs. With that as a backdrop, CIOs along with IT experts must consider their organizations' business needs, and the potential threats that system downtime and potential data losses pose to the success of their organizations. After thoroughly analyzing business needs and potential threats to the data asset, the next activity is to apply the appropriate technologies to ensure that efficient backup schemes and proper protection are in place for the organization's data asset. Once a solution is in place, the organization must rigorously adhere to best practices that support the solution.

What this white paper provides

This white paper provides:

- An examination of the critical need for data protection in today's information-rich organizations
- Key business and technology drivers that help determine the functional requirements needed from various data protection technologies to meet a set of business and technology needs
- High-level overview of various data protection technologies and the optimal positioning of each

Intended audience

This discussion should first be of interest to CIOs, needing technical knowledge to help make informed decisions about backing up and protecting their organizations' data assets; also, to those IT professionals, and systems administrators seeking the same knowledge.

Introduction

Data protection concept

The concept of data protection once prevailed as a relatively straightforward practice. Ever since the commercial introduction of magnetic-data storage devices, in 1947, users of this technology became aware of the vulnerability of the data stored on these devices. Therefore, the introduction of data backup and recovery was quickly introduced as a means to protect the investment made in data stored on this medium. Originally, the goal was simply to create a second copy of a primary data set, storing it on a less expensive medium than the primary medium. This second copy was retained and generally only accessed in the event of a primary disk failure, where it would be copied onto a new or repaired disk.

Data protection concept evolution

As the computing industry grew, however, a number of pressures rendered inadequate the original data protection model. Through the years, some of the trends forcing the model to evolve technologically are the following:

- Decentralized computing: Over the past fifty years, the computer industry has been the subject of several shifts in computing focus. In the early years, resource centralization was standard and clients were given an interface to send commands and receive responses from the central computer. Advances in technology led to the distribution of computing power to a decentralized model that enabled multiple computers to work together by distributing tasks between client and server machines to achieve a more robust environment. This trend also led to the distribution of data to the various computers, providing optimal performance in a networked environment. That meant system administrators who were responsible for data protection now had to manage data stored in a variety of places, rather than in a single centralized repository. In addition, decentralization put considerable upward pressure on IT budgets due to administrative costs.
- The cost of data: As data volumes grew, so did the organization's investment in the maintenance of that data. According to a recent IDC industry study, the cost to recreate just twenty megabytes of lost data, if even possible, is estimated at:
  - 19 days @ $18,700 for a commercial department
  - 21 days @ $21,000 for an accounting service
  - 42 days @ $107,000 for a research and development firm

Continued on next page

***Introduction,** continued*

Data protection concept evolution, continued

- Data growth rates: The rate at which organizations generate and store data has been tremendous over the last decade, and data volumes are projected to continue to grow at a rate of 80% annually through 2005. In the current economy, it is safe to say that IT budgets on average have not kept pace with the data growth rate. This puts continuous strain on the organization's data storage and management infrastructure, particularly in maintaining adequate protection and recoverability of data assets.
- System uptime requirements: In the past, organizations scheduled data backups to occur during system downtimes to minimize the impact on users and other processes that require computing resources, during the typical nine to five business day. In recent years, system administrators have felt increasing pressure to reduce or eliminate system downtime to enable production on a 7 x 24 x 365 basis.
- Network resource contention: In a decentralized computing environment, data backup and restore operations can contend for the same network resources as production users. This can lead to bandwidth saturation, potentially crippling production during backup operations.
- Heterogeneous computing environments: Along with the decentralization of computing resources came the introduction of heterogeneous computing environments. In today's enterprise environments, it is common to find multiple computer platforms and operating systems, all connected by a common network topology.
- Data complexities: Data types such as databases, spreadsheet compound documents, web sites, and collaboration systems introduce elements of complexity that were never imagined in the early days of data protection. With these data types, the relationship among multiple files is critical to the integrity of the entire data set; and if that relationship is not preserved when the data is backed up, the entire data set could be rendered worthless. This forces the need for data-cognizant backup applications that help ensure the integrity of each data source either backed up or restored, and for synchronization.
- Advanced storage architectures: The explosion in data generation and storage has forced a paradigm shift in storage architecture design. This trend has led to the introduction of advanced storage architectures such as Storage Area Networks (SANs) and Network Attached Storage (NAS). These architectures separate application servers and file servers from storage subsystems:
  - To improve access to data
  - To improve administration efficiency via consolidation
  - To isolate backup traffic from the Local Area Network (LAN)

  Although these architectures offer substantial benefits to the data backup, recovery, and protection processes, they also impose new challenges that will be discussed later.

Continued on next page

Current need for data protection, *continued*

The goal of data protection

To successfully design and implement an open system data protection solution in today's enterprise, one needs to fully analyze the business needs of the organization, and map the correct suite of technologies to meet those needs. It is imperative to start at the highest level in the organization to ensure that each component of a data protection strategy fully supports the organization's mission. For example, if a data-driven organization has as part of its mission, to become the dominant world-class leader in its industry, how would that goal affect decisions regarding protecting its data asset?

Responsibility to protect data assets

CIOs have the foremost responsibility to ensure that their organizations' data assets are recoverable and protected. Therefore, it is essential that CIOs understand the spectrum of potential threats to their data assets, and then become informed about the technologies used to thwart those potential threats. That understanding will help CIOs enter into intelligent conversation with IT staff when discussing the appropriate technology choices to protect data assets.

Topics in this paper

The storage industry currently takes a highly compartmentalized view of data protection technologies, placing them in narrowly defined categories such as:

- Backup
- Redundancy
- High availability
- Continuance
- Disaster recovery

While these categories have merit to define and differentiate technologies and product offerings, it is equally important to understand the relationship among the categories, and the technologies they represent. The relationship starts with one of the organization's most valued assets, its data. This paper presents a discussion that seeks to shed light on those relationships by addressing the following topics:

- Data protection
- Data protection assessments
- Mapping technology to those assessed needs
- Terms and definitions

Data Protection

Current need for data protection

Data backup and recovery systems

In recent years, several trends have caused IT professionals to become more diligent in designing and maintaining data backup and recovery systems. Some major factors to consider with regard to justifying and designing an appropriate data protection solution include:

- Growth in stored data
- Cost of downtime and data unavailability
- Cost of data management
- Data as a core business asset

Growth in stored data

Data storage has grown rapidly in recent years. Conservative estimates from the International Data Corporation (IDC) show data expanding at approximately 80% per year, while other industry analysts place the growth rate closer to 100% per year, even in the current struggling economy. There are many data types driving this growth, and at the enterprise level these include:

- Databases
- Email
- Multimedia

Databases: Data within database management systems comprise approximately 55% of all data on disk subsystems across enterprise, midrange, and distributed computing platforms. The following database integrated applications contribute to the generation and storage of Database Management Systems (DBMS) resident data:

- Enterprise Resource Planning (ERP)
- Supply Chain Management
- E-procurement
- Content Management
- Data Mining
- Customer Relationship Management (CRM)
- Electronic Document Management (EDM)

This statistic is meaningful from a backup and recovery perspective because database-resident data can be challenging to successfully backup and restore in production environments.

Continued on next page

Current need for data protection, *continued*

Growth in stored data, continued

Email: In the last five years, email activity has risen from a novelty status to an essential core-business communications tool. Approximately 40% of Americans currently utilize email communications at home or work, and the worldwide mailbox count is estimated to exceed one billion. Virtually all-large organizations have become dependant upon email as a vital, core communications tool. In addition, as bandwidth has become cheaper and more widely available, there has been an increase in the volume and size of email attachments.

In addition, in 1998 the Securities Exchange Act and the Internal Revenue Service codes were amended, requiring organizations with ties to the securities industry to maintain all electronic correspondence between employees and clients. Combined, these factors led email communications to drive approximately 900 terabytes of storage in the United States alone, in the year 2000.

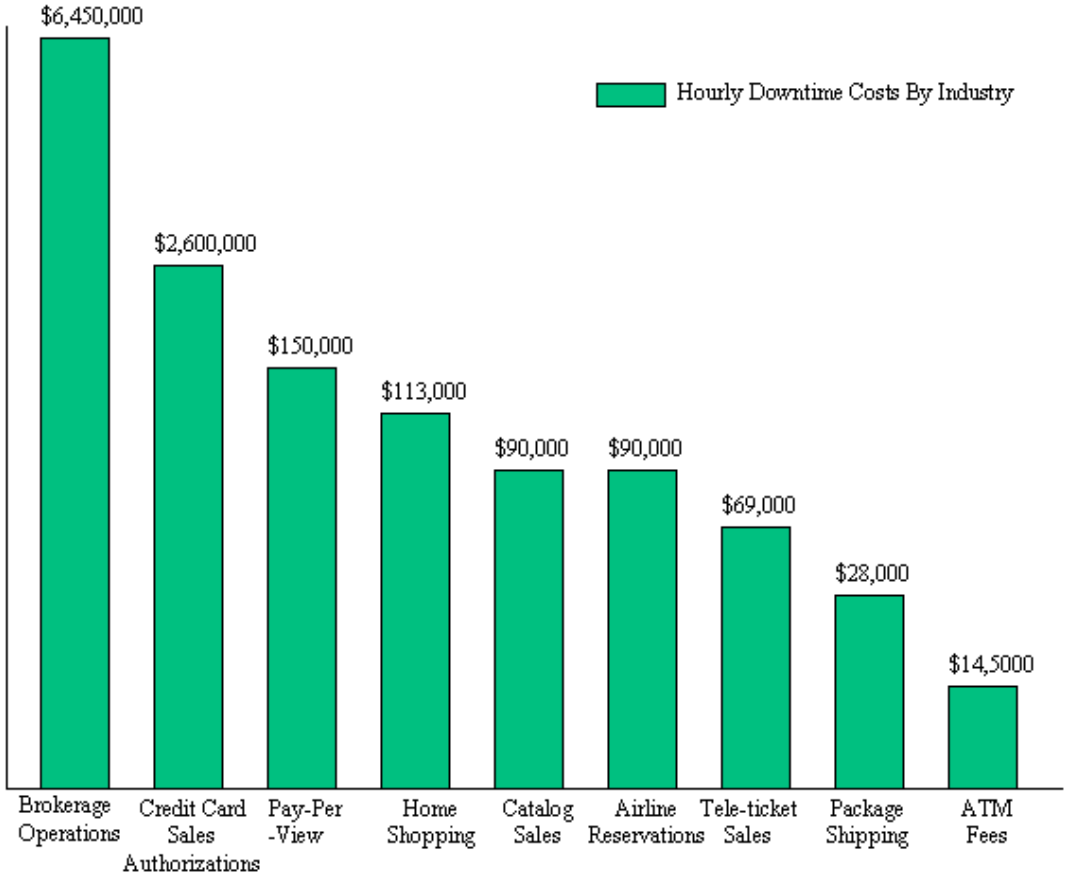
Multimedia: Applications used for audio, video, and graphics creation and manipulation generate large files that consume tremendous amounts of storage. Marketing communications groups increasingly leverage these technologies as part of their overall corporate branding efforts.

Continued on next page

Current need for data protection, *continued*

Cost of downtime and data unavailability

Other factors to consider are the high cost of system downtime and data unavailability. Organizations leverage their data assets in ever-increasing proportions, leaving little time or tolerance for system downtime or critical data not being available. The Fibre Channel Industry Association chart shown below depicts the high costs to various industries, resulting from downtime and data unavailability. These figures indicate the average cost for the first hour of downtime in each given industry. With each passing hour, however, it is likely that this figure actually rises and the likelihood of a long-term impact on the organization is higher. Data unavailability could become a disaster at some point, with the potential for terminal results for the organization.



What the numbers suggest

As these exorbitant costs suggest, it is vitally important to attain a clear understanding of the costs associated with downtime in one's organization, as a starting point to determine a viable data protection strategy. These figures also suggest that it is wise to plan and configure a backup solution that is not only tuned to meet backup performance and backup window objectives, but also tuned to meet the organization's acceptable downtime-recovery window. Because day-to-day administrative pressures focus on meeting the ever-shrinking backup window, it is essential to remain mindful of the original purpose for backing up data, which is the ability to restore usable data in a timely manner.

Continued on next page

Current need for data protection, *continued*

Cost of data management

As the amount of data generated and stored by organizations continues to grow at a staggering rate, so does the cost associated with managing that growth. Many sources estimate that the cost of managing storage in the enterprise, including technology and human resources, is anywhere from three to ten times the cost of the storage itself, depending upon the degree of storage centralization and consolidation within the environment. Depending upon the strategy and technology deployed in a given environment, administration of data protection systems can represent a substantial portion of an organization's data management costs.

Data as a core business asset

Today organizations are increasingly dependent upon their data as a core business asset. They increasingly find themselves in dire straights if their data is unavailable or lost. Without key data, many organizations would quite likely cease to exist. Consider some facts pertaining to the potential impact of data losses on businesses:

- On average, companies value 100 megabytes of data at more than $1 million dollars according to a recent study by Ontrack Data International. With hard drive prices dropping to less than one penny per megabyte for low-end disk drive technology, disk cost is not a valid measure of the value of the data stored on disks.
- Estimate show that 1 out of 500 data centers will have a severe disaster each year.
- A recent survey by IBM indicates that only eight percent of Internet companies are prepared for a severe system disaster.
- According to Ontrack Data International, computer crimes cost firms that detect and verify incidents of computer crime, between $145 million and $730 million each year.
- A company that experiences a computer outage, lasting more than 10 days will never fully recover financially; 50% percent of those will be out of business within five years.

These statistics underscore the importance of data to today's industries. They also validate the important need for comprehensive data protection strategies.

Data Protection Assessment

Know your data

Data types and uses

A basic data protection assessment should include a review of all pertinent data types, along with the availability needs for each type of data. An assessment of the data types should answer the following questions:

- What would be the cost to the organization to recreate each type of data? The higher the cost, the more extensive the data-protection strategy should be. For data that can be recreated without extensive costs or delays, lower levels of protection should be adequate.
- What is the cost to the organization for each hour that a data source is unavailable? Considerations include lost productivity, lost revenue, and the possible loss of customers if the period of unavailability reaches a certain threshold. This information is used to determine the organization's Recovery Time Objective (RTO).
- What is the projected business impact if data is permanently lost and cannot be recreated? This information is critical when assessing the requirement for data redundancy and off-site data vaulting.
- In the event of a system failure, how many transactions can the organization afford to lose? In other words, how frequently must backups be performed? This information is used to determine the organization's Recovery Point Objective (RPO).
- How much regularly planned system downtime or slowdown time can the production environment tolerate for backup activity? In other words, what is the backup window?
- How much administrator time is currently allocated to backup activities per terabyte of data? Can this ratio be sustained at your current data growth rate? Consider technologies that would automate and centralize backups to improve storage administration efficiency. This should be done with a clear understanding of the potential Return on Investment (ROI) for each technology. The key is to assess those that would yield the most benefit to the organization.
- What is the retention period for each data type? Retention requirements impact many infrastructure design decisions, such as the storage medium, rotation policies, library size, vaulting service selection, and backup data redundancy.
- Is there a requirement to store a copy of the data off-site? Off-site archival of backup media can influence several decisions, including backup software, vaulting service, and storage media.
- Is this data currently mirrored or replicated within the environment or at a disaster recovery site? Duplicate copies of the produced data enable the consideration of such activities as off-host backup, using a third copy of data, or off-site backup performed at a replica site.

Continued on next page

Know your data, *continued*

Data types and uses, continued

- How much data of each type is currently stored within the organization?
- What is the rate of growth for each data type?
- What is the configuration of the server(s) that hosts this data, if applicable? Include make/model, number of processors, operating system and level, and a list of applications hosted by each system. This platform information can impact the data protection infrastructure.
- What are the data types and amounts stored on user workstations vs. the network and application servers? Include a list of each operating system, along with an approximate number of instances of each in the environment.
- Is workstation-resident data currently protected by backup software? If so, how much administrative time is dedicated to workstation backup?
- How long would it take to rebuild a workstation in the event of a failure, or if it were stolen?
- Are there legal or regulatory requirements for data retention?

What's the next step?

Now that the data types, storage locations, amounts, and availability requirements are analyzed, the next step is to determine data recovery times, recovery points, and backup windows to effectively backup and recover the data.

Know your data backup and recovery needs

Ability to backup and recover data

After gaining an understanding of the relationship between the data asset and the business needs, it is necessary to determine the backup infrastructure's ability to backup and recover data. Setting recovery time, recovery point, and backup window objectives will expose the backup infrastructure's ability to meet the data backup and recovery needs of the organization.

Recovery time objective

Recovery Time Objective (RTO) is the measurement of how long it takes to return an identified data source to accessibility following a named data interruption. For example, if an organization's RTO is to have a certain data set available within two hours, following a disk-subsystem failure, that objective should drive the infrastructure design to protect that data set. When the objective changes, for example, to a two-minute RTO, the data protection infrastructure must be configured to meet that requirement.

Different RTOs may need to be developed for different types of data, and potential data access interruptions. For example, an organization might determine that it needs to be able to recover its email data within one hour of a common system failure. The organization may determine that it can afford a greater amount of time to recover that same email data at a remote site, in the case of a server-room disaster. In that case, the level of investment required to yield a two-hour recovery time for email data may not be justified, given the unlikely possibility of a server-room disaster.

Recovery point objective

Recovery Point Objective (RPO) is the measurement of the point in time that a data source is restored, following an event that caused data loss. For example, if an organization experiences database corruption, and uses only a traditional tape backup, its point of recovery for that corrupt database would be from the last backup, which could be from the previous night. While last night's copy was once a commonly acceptable RPO, today most organizations strive to reduce the amount of data exposed to risk during the production day. You can reduce RPO by implementing technologies that allow live, frequent, and strategically timed recovery points that minimally impact production systems.

As with RTOs, it will be necessary to establish different RPOs for each data type within the organization. For example, business critical data with the highest refresh rates should be given a shorter RPO, while less critical data with lower refresh rates can be given a longer RPO.

Continued on next page

Know your data backup and recovery needs, *continued*

Backup window

A backup window is the time frame that storage administrators can safely shutdown applications, drop user connections, and make the system safe for backup operations. Most administrators refer to this sarcastically as something that existed in "the good old days." While many administrators have seen their backup window give way entirely to production demands, many others still struggle to make the most of the limited backup window they still have. In either case, it is important to design a solution that meets the organization's backup window requirements, but without losing sight of recoverability.

Backup systems are generally used nightly, so challenges to meet backup windows could take priority over the challenges to meet RTOs. Given that, one must keep in mind that the sole purpose of the backup solution is to recover data. While backup windows are important, the solution should not be overly tuned to meet backup window requirements at the expense of data recoverability. Therefore, any decision to implement technologies to better meet backup window requirements must be weighed against the potential impact that the technology will have on the recovery time for the data source involved. Note the following examples:

- It may be possible to reduce the backup of a large-file server by 50 to 75 percent. This is made possible by implementing block-level backup as a replacement for file-level backup, but if file-level restorations are commonly needed in production, where performance is critical, the net impact on the environment could be undesirable.
- Use of data interleaving *(also known as multiplexing)*, which combines multiple backup data streams into a single stream to help push tape drives at rates for best performance. While this tunes the environment for backup performance, the additional data that must be read by the tape heads during restore, and the system processing time for sorting the data can have an extremely negative impact on restore performance.

What's the next step?

Now that the data protection assessment is completed, the next step is to choose the right technologies to augment the organization's storage infrastructure.

Mapping Technologies to Business Needs

Needs based application of technology

Upon completion of a comprehensive data backup and protection assessment, the business needs can be mapped to available technologies that would ultimately render an optimized solution for the organization.

It is important to keep in mind that data backup and protection is not a one-size-fits-all model, because the application of technologies must be based upon the specific and sometimes specialized needs of an organization. In no particular order, some current data backup, recovery, protection, and management technologies available in today's marketplace are the following:

- Block-level backup
- Removable media management
- Open file protection
- LAN-free backup
- Serverless or server-free backup
- Point-in-time off-host backup *(mirroring)*
- Snapshots
- Disk staging
- Data replication
- Centralized backup administration
- Encryption
- Bare metal restore
- Hierarchical storage management (HSM)
- Workstation backup

File-level and block-level backups

How are file-level backups done?

The host processor, during a file-level backup makes numerous I/O and system calls to locate and open the header of each file for backup. This data is commonly fragmented across a set of disks, requiring a high frequency of scattered disk I/Os to accomplish the backup. This can cause a number of undesirable effects in environments with large numbers of small files, including:

- Server slowdown due to high processor utilization
- Poor tape performance, resulting from frequent rewinds to establish tape repositioning *(known as "shoe shining" due to the back and forth motion of the tape on the tape heads)*
- Accelerated wear and tear on tape heads and tape resulting from shoe shining

Note: File-level backup is the default process.

How are block-level backups done?

The block-level backup process creates a bitmap image of the file system's storage blocks. Once the image is established, the storage blocks represented by this image are moved to a secondary storage device, without further referencing this file system. As data changes occur during the backup process, they are written to the production data volume, while the blocks that are to be backed up are copied to a cache area on the disk drive. When the backup application finds a changed block, it refers to the cache area for the accurate point-in-time data. Upon completion of the backup, the cache is cleared.

When applying incremental block-level backup technology, additional optimization can be achieved in performance and in tape utilization. The difference between incremental-file backup and block-level backup can be illustrated as follows: If you create a document using word processing software of 1 Megabyte (MB) in size, this might represent approximately 170 data blocks of 6 Kilobyte (Kb) each. Then if you change one word in that document, a file-level incremental backup would copy the entire 1 MB file, conversely a block-level incremental backup would only copy the 6 Kb change.

When to use block-level backup

Environments with large volumes of small to medium sized files are good candidates for block-level backup technologies, as backups can be performed in shorter time periods with less processing impact on critical application and file servers. The main drawback to block-level backup is that while it tunes an environment for backup performance, it tends to detune for restore performance on file-level restorations. In other words, there could be a gain in backup performance, but at the expense of recovery time.

Removable media management

Tape as removable media

Many data protection architectures include removable media, such as tape as the main storage medium for backing up data. Current tape library technology allows for the import and export of tapes, while indices to the data on the removed tapes can remain online. It is desirable to remove and track these tapes for the following reasons:

- Optimal tape library sizing: It makes sense to implement a library that will accommodate regularly accessed volumes with room for growth, but it is overkill to store an entire archive of tapes that is either unlikely to be accessed, or the RTO for the stored data set allows for manual retrieval of the tapes.
- Many organizations, as a practice, create a second copy of each tape to protect against failures. This is particularly common and wise in Hierarchical Storage Management (HSM) and archive applications, where the data on tape is not a backup copy, but in many cases, the *only* copy.
- Rotation of tapes to off-site storage for disaster recovery purposes.

Managing and tracking tapes

Software-based vaulting utilities are available to manage and track tapes as they move through their life cycle from library to shelf, and to off-site storage for disaster recovery purposes. These products generally provide the ability to generate a report on the life cycle status of each tape; also, automatically export tapes from the library due for export, and generate a pick list for items to be returned from off-site storage. This technology automates a once tediously manual process that was highly error prone.

When to use vaulting technology

As data volumes grow and backup operations expand, many organizations experience challenges, and spend innumerable hours of a valuable system administrator's time, manually tracking backup tapes. In such environments, removable media vaulting technology can greatly simplify the administration of tape import, export, and life cycle management.

Open file backup

Open file backup software

There are software technologies available for managing open files during scheduled backups. Essentially, this software uses a caching area on disk to capture changes that occur during the time of backup. This open-file backup technology is integrated with many leading backup software products and can be used to provide a solid solution for backing up files in high-transaction environments, or environments with a small or non-existent backup window.

How are open files backed up?

With the increasing widespread movement toward 7 x 24 x 365 operations, a common source of backup and recovery frustration is that certain files are frequently in use during backups, which can compromise the integrity of the backup for several reasons:

- A file can be skipped altogether if the backup software is configured or designed to do so
- A file could be unavailable during the backup process, given that the file would be essentially open while being backed up
- The backup might be forced on files in use

When a forced-file backup occurs, any updates to that file that happen during the time of the backup might be captured by the backup software, while other changes might be in areas of the file that have already been read for backup, resulting in inconsistent or corrupt data. In this case, the file contained in the backup might be completely unusable.

When to use open file protection technology

Any organization that performs backups on file servers or workstations that have any potential for files to be in use during the backup process could benefit from this technology.

LAN-free backup

LAN-based backups

LAN-based backups transfer backup data across the network in order to centralize tape administration and leverage network resources. See Figure 1. This is a good method, if the backup activity can be contained within an acceptable backup window, and the LAN is not used by other applications during backups to prevent network saturation.

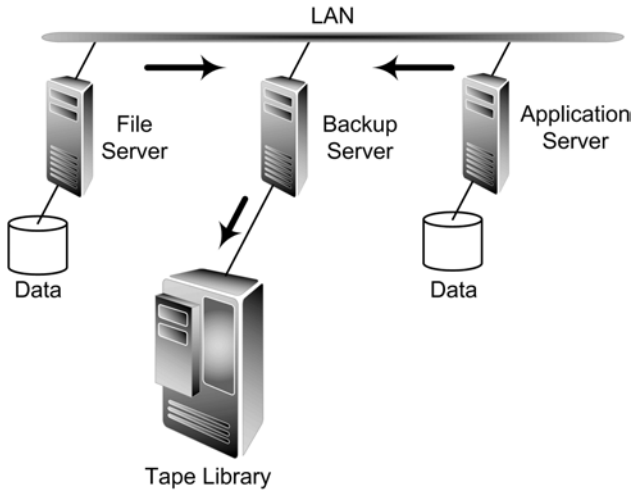


Figure 1: LAN -Based Backup

A case for LAN-free backup

A major challenge with LAN-based backups is that many organizations are experiencing shrinking or disappearing backup windows. The business requirements in these organizations are such that the production network is active 7 x 24 x 365, and cannot tolerate the saturation introduced by backup traffic. That scenario calls for other methods of backing up data, where the messaging LAN is not used for backups—advent LAN-free backup.

Continued on next page

***LAN-free backup,** continued*

LAN-free dedicated TCP/IP backup network

There are a couple of prevalent strategies to minimize traffic on the LAN while backing up data. One strategy is to implement an additional TCP/IP network, which is dedicated to the task of transferring backup data to shared backup storage subsystems. See Figure 2. Essentially, additional dedicated network interface cards and cables are added to each server, and any necessary additional network components, such as switches, hubs, routers, etc. Most backup software products can operate transparently in this type of environment.

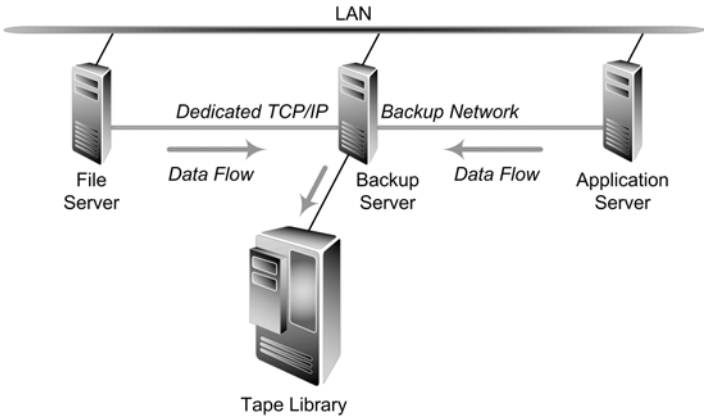


Figure 2: LAN-Free Dedicated TCP/IP Backup Network

Continued on next page

LAN-free backup, *continued*

LAN-free backup over a SAN

Another strategy for minimizing backup-related traffic on the LAN is to implement a Storage Area Network (SAN). SANs house both the primary and secondary storage subsystems, and during backup operations transport the data from the primary to the secondary storage subsystem without impacting the LAN. See Figure 3.

In addition to isolating backup traffic to the SAN, preventing network saturation, this strategy also generally provides higher data-transfer rates than traditional LAN topologies. That serves to reduce the time spent backing up data, leaving host systems available when needed. By implementing a SAN, the environment is also well prepared for serverless *(also called server-free)* backup, which is now supported by most major backup applications.

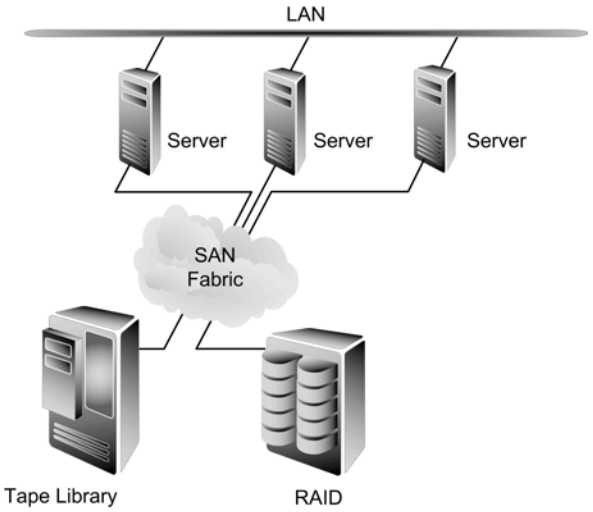


Figure 3: LAN-Free SAN Backup Network

LAN-free backup challenges

With an open system such as a SAN, many different file systems, volume management and disk management formats, and software demand that security be considered and implemented within the SAN to keep servers from gaining unauthorized access or accidental access to certain data resources. Given that, not only must the organization assess its data protection needs for potential disasters and external threats, but also the same must be done to contain the technology within the SAN.

When to use LAN-free backup technology

When network saturation causes unacceptable delays in backup operations, LAN-free backup might be a desirable option. In addition, in environments where incremental traffic created by backup operations causes sluggish response to production users of network resources, this strategy can be utilized to improve productivity.

Serverless (server-free) backups

Concept of serverless backup

The concept of serverless backup has been touted as an up and coming technology by vendors for several years, but in reality it has trickled into the marketplace rather than taking it by storm. The allure of serverless backup is that it allows backups to occur at any time, without taxing the production server with extensive file system thrashing or I/O processing. This would allow organizations with no backup window to achieve point-in-time backups without impacting production, and without requiring extensive additional primary storage space.

How does serverless backup work?

With serverless backup, the system hosting a data source takes a snapshot of the file system or a bitmap image of the block status on the storage subsystem, and assigns the I/O load associated with data movement to another server or SAN resource such as a router, switch, tape drive, or a dedicated server. In this approach, the data host is taxed for only a brief moment while it compiles a list of the backup data. This corresponding list of storage blocks is then transferred to the data mover, where that unit uses the Small Computer System Interface (SCSI) extended copy command to copy data from the primary to the secondary storage subsystem. During the backup, host resources are unaffected, freeing the host to respond effectively to client requests for either applications or file resources. As data updates occur, during the course of a backup, they are written to a temporary holding area until the backup is completed.

As the data mover successfully migrates data from the primary to secondary storage subsystem, it relays metadata associated with this backup to the backup server, and the pending updates are applied to the primary data set. See Figure 4.

On a restore, the process essentially works in the same way, with the data movement going from the secondary to the primary storage subsystem.

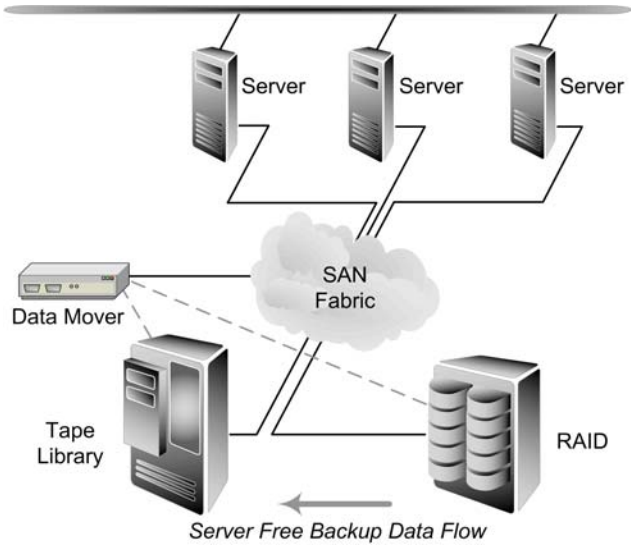


Figure 4: Server-Free Backup

Continued on next page

Serverless (server-free) backups, *continued*

When to use serverless backup technology

Serverless backup is a relatively new and sophisticated technology. It should be considered for high volume On-Line Transaction Processing (OLTP) environments with large data volumes and unmanageable backup windows, as a means of preserving host system performance during backup operations. Serverless backup should also be considered for environments where a backup window is not available, or the volume of data to be backed up makes it impossible to complete a backup within an acceptable backup window. Serverless backup is still very complex technology; it should only be used when more established backup methods would not meet the organization's backup window requirements.

Point-in-time off-host backup

How does disk mirroring technology work?

Many storage infrastructures include component redundancy in the form of disk mirroring, using RAID (Random Array of Independent Disks) 1 to improve the reliability and availability of primary storage subsystems. Given the declining costs of disk, increasing numbers of organizations are taking advantage of mirroring technology to augment their backup capabilities. By deploying an additional mirror, this third copy of the data can be separated from the primary and the first mirror for backup operations. This provides a point-in-time copy of data, which is separate from applications and user data. A separate server can back up this copy without adversely impacting production on the application server. See Figure 5.

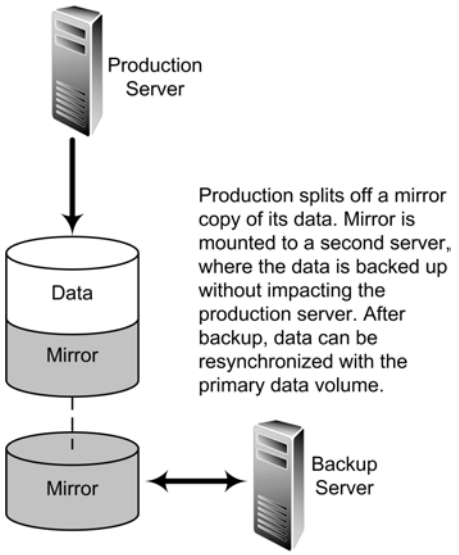


Figure 5: Mirroring

Third mirror management

Third mirror management can be performed by either hardware or software-based utilities with similar benefits to backup operations. One advantage of using software-based volume administration to manage the point-in-time copy of the data is that inexpensive storage media can be used for third mirrors. This allows organizations to invest in cutting edge RAID technology for production storage, while utilizing less expensive storage subsystems, or even repurposing legacy storage for their third mirrors, where performance and availability are not as critical as on production storage.

When to use mirroring technology

Off-host backup provides a method for delivering quality, predictable, point-in-time data protection, using standard backup technologies. Given the declining cost per megabyte of disk storage, this approach is gaining popularity as a way of performing backups without hindering production operations in environments with shrinking or non-existent backup windows.

File system data snapshots

How does snapshot technology work?

Snapshot technology provides a means of creating parallel, read-only file systems that point to a set of data intermingled with live-production data. Creating file system snapshots take only seconds with minimal impact on the system. These snapshots are stored as small files on the live file system. The data that exists at the time of the snapshot is protected from being overwritten on the physical disk so that it may be referenced from the snapshots. This enables consistent, static access to files at an identified point-in-time, which offers tremendous benefit to both backup and recovery processes.

Snapshot data edits, additions, and disk-space requirements

Data edits and additions are written to a new area on the disk, which means that snapshots do not require nearly the incremental disk space required for point-in-time data copies *(split mirrors)*, but generally some incremental disk space is required. The disk-space requirement is dependent upon how long the snapshots are kept, and the refresh rate of the data. It is important to manage and cycle the snapshots so that the unneeded disk space can be released and made available to the live file system. See Figure 6.

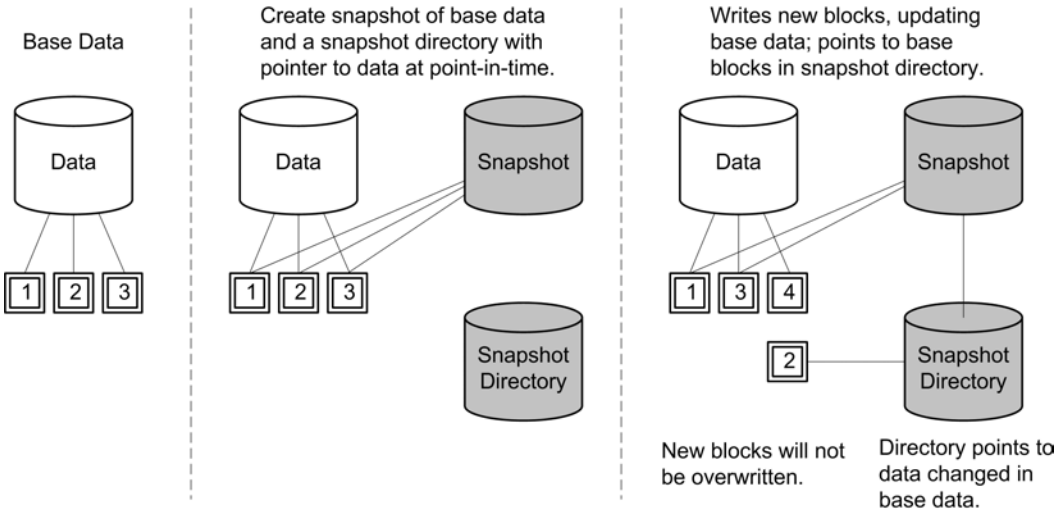


Figure 6: File System Data Snapshots

Continued on next page

File system data snapshots, *continued*

Snapshots used as a part of the backup process	Snapshots of data can be taken either to create a consistent point of quick rollback for cases of inadvertent changes, deletions, or corruption; also, to establish a solid point-in-time reference to a live data source to assist backup operations. When snapshots are used as part of backup, a snapshot of the data is taken before the backup process begins. Then the data host mounts the read-only snapshot file system for backup purposes, while continuing its production use of the live file system.
Snapshots used for data recovery	For recovery, snapshot file systems may be referenced to restore files that have been corrupted or inadvertently deleted. In many environments, snapshot technology is used for up to 90% of file recovery, rather than retrieving the file from tape or other secondary media. This approach to recovery greatly improves performance and ease of administration, and serves to complement traditional backup and recovery technologies.
When to use snapshot technology	Snapshot technology brings consistency to backups offered by a point-in-time file system that cannot change during backup operations. Snapshots offer a great deal of benefit to data protection operations at lower price points. On file servers and NAS platforms where this technology is available, it is highly beneficial and is worth considering in virtually any environment. For complex relational data types such as databases, it is generally preferred to use the DBMS tools to establish point-in-time rollback capabilities so that pending transactions can be fully applied. This approach will provide a solid rollback point for recovery.

Disk staging

How does disk staging work?

Disk staging, sometimes called data staging or disk-to-disk-to-tape (DDT) backup, allows data backups to tape in a two-step process:

1. Data is copied across the LAN from a data host to a disk volume hosted by a backup server.
2. When certain thresholds are met or the data copy is complete, the data can be written to tape by the backup server. When the data is ultimately written to tape, it is written over a high-speed bus, such as SCSI or fibre channel.

Data staging can also provide a foundation for performing Synthetic Full™ backups, which allow organizations to effectively perform a single-full backup, followed by ongoing incremental backups without ever having to perform another full backup; merging the full backup with the incremental backups happens behind the scene, resulting in an updated full backup. This allows an organization to perform a full backup, while only causing the network traffic and application impact of an incremental backup. See Figure 7.

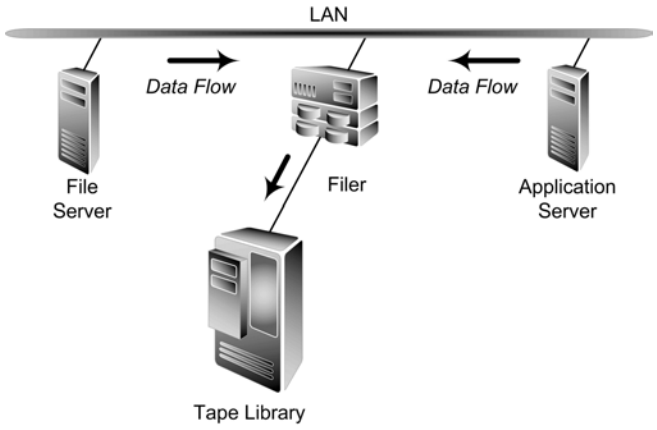


Figure 7: Disk Staging Architecture

Challenges with data staging

Although this approach may appear to be a panacea, there are drawbacks that some environments cannot tolerate. For example, depending upon how the backup application performs the two-step backup process: disk and volume merging, there will be variable amounts of disk space used in this process. In some cases, this could compound the amount of disk space needed for a given data repository, making the ROI impossible for certain applications.

Other side effects may include increased server utilization during a volume consolidation process, and excessive wear on tape drive and media from cycles spent on data consolidation.

Continued on next page

Disk staging, *continued*

When to use disk staging technology	Disk staging can be used in environments with inadequate network bandwidth to provide sustained throughput at a level sufficient for streaming a tape device. In such environments, the data can be staged on disk prior to transferring to tape, which allows for optimal tape drive utilization.

Data replication

Data replication

Data replication technology provides the ability to create a secondary copy of data, generally for the primary purpose of disaster recovery. Software-based and hardware-based replication technologies each have strengths and weaknesses. By using hardware-based technology for replication, the data transfer can be performed without impacting the application or file servers. On the other hand, software-based technology better integrates with the file system and applications, allowing better data consistency, which is particularly critical for database applications. In addition, on modern server architectures, the additional host cycles for software-based replication are generally considered nominal.

How does data replication technology work?

Data replication technologies are diverse technically; and at a high level, they fall into these two categories: scorecard or write duplication.

1. In scorecard replication, a baseline file system or disk-block bitmap is created, and periodically monitored for changes. These changes are then applied to the replica data.
2. In write duplication, disk writes are intercepted by the replication software, and applied at the replica site in the same order as written to the primary site. Write duplication happens synchronously, asynchronously, or in some cases, semi-synchronously.
   - Synchronous data replication applies each write to the replica system while the primary host waits for verification. In environments with bandwidth constraints, this can have performance implications on production storage.
   - Asynchronous data replication places each write into a queue, so that primary storage performance is not affected by Wide Area Network (WAN) latency.
   - Semi-synchronous data replication allows the system to operate in synchronous mode until certain performance thresholds are reached, and then automatically switches to asynchronous mode.

See Figure 8.

Continued on next page

Data replication, *continued*

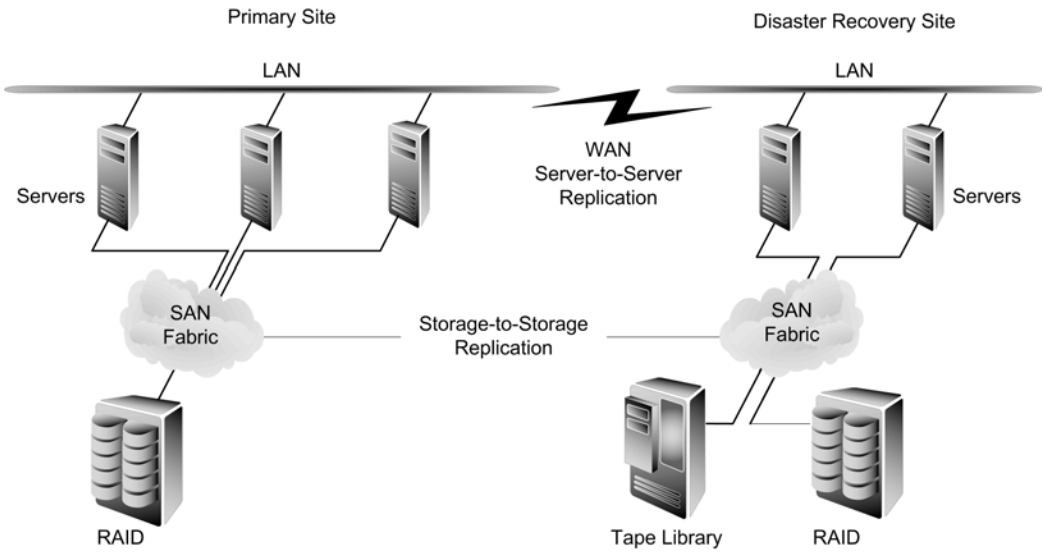


Figure 8: Data Replication Architecture

Data replication and data protection

While data replication provides benefits generally associated with disaster recovery, the replica data can be leveraged for general data protection purposes as well. Some organizations perform their regular backups on the replica data, which yields the additional benefit of off-host backup. This approach can also be used to centralize the backup operations of multiple facilities.

When to use replication technology

Replication technologies are largely implemented as components of a business continuance, disaster recovery, or high availability strategy. When these technologies are present, they can be leveraged for backup purposes by providing faster restore times from storage-system failures, or they can provide off-host backup capabilities.

Centralized backup administration

Centralized data protection

In an enterprise-class backup environment, one important consideration is to maximize administrative efficiency, and minimize the human-resource time devoted to planning, scheduling, and administrating backup systems. One common characteristic of large enterprise data-storage environments is data being housed at multiple physical locations. These locations could be disbursed across a campus, region, country, or the world, but the implication is the same. Steps must be taken to incorporate each data repository into a centralized data protection scheme, or the remote data will require inefficient, fragmented backup administration.

How does backup centralization administration work?

The backups can be physically centralized by moving all the remote data streams to a common backup infrastructure, or virtually centralized by implementing sophisticated hierarchical software technology for backup administration.

- Physical centralization requires adequate bandwidth to facilitate the data throughput necessary to meet the defined backup window.
- Virtual centralization requires software that provides a global view of the data protection infrastructure from a single access point, and is capable of managing the distributed backup products throughout the enterprise. In some cases, this approach offers the ability to assign varying degrees of control to backup administrators and operators, based upon minimal requirements for localized control in their own environment.

When to use centralized backup administration technology

Backup vendors offer varying degrees of centralization to accompany their products, therefore, should be considered at some level in most organizations. As the backup requirements expand and become more complex, centralized administration allows the backup systems to scale with the requirements, without the linear expansion of IT human resources to facilitate the expansion.

Encryption

Encryption security

Backup systems have a reputation for creating data sets that are highly susceptible to unwanted access. This data is particularly vulnerable at three points:

1. During transfers across the network from a backup client to the backup server
2. During stores to tape in a standard format
3. During transfers across the network from the tape back to the client

For environments that are security sensitive, it may be advisable to implement encryption technology to prevent data from being intercepted during transfers across a network from a backup client to the backup server. It may also be advisable to encrypt the data as it is written to tape, so that as the tape is placed on a shelf or transferred off-site it does not fall prey to malicious tampering. See Figure 9.

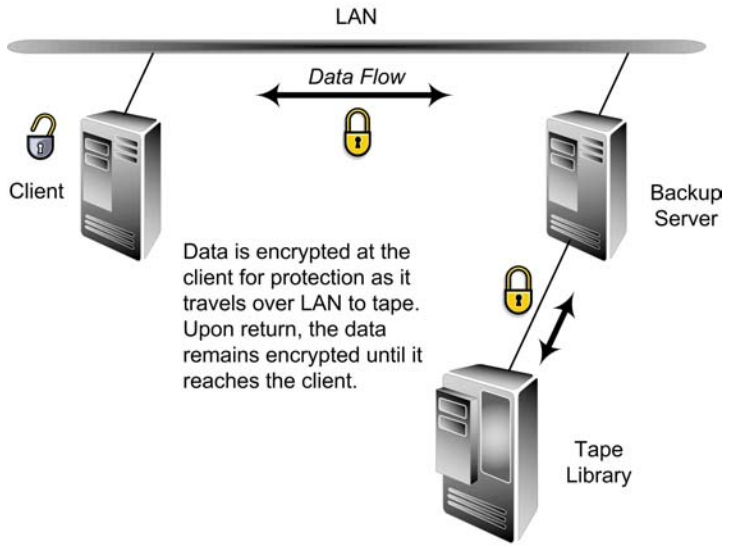


Figure 9: Encryption

Challenges with encryption technology

The trade-off in considering this technology is that it can add overhead from incremental-system processing during the backup process. It can also make the backup data unreadable by standard system utilities, forcing the administrator to use the original backup software to read the tapes. This can be undesirable as an organization considers changing backup software products, or attempts to consolidate multiple backup products within a decentralized environment. This is compared to non-encrypted backup data, which is often written in standard formats for data portability purposes.

When to use encryption technology

Organizations with a high need to secured data should consider encryption technology as an additional measure for protecting data assets.

Bare metal restore

Bare metal restore automates disk restoration

In the event of a failure that renders a host system unbootable, the time to restore the system to full functionality, along with all its hosted data, can be extensive. Before recovering the data, the operating system must be restored, patched, and configured, and all applications must be installed and configured as well. As an alternative, bare metal restore technology can be deployed to reduce significantly the human intervention of restoring a system from a raw drive to a fully restored production system. The technology is generally designed to reduce the number of steps necessary to restore a server to a usable state, as well as shorten the time needed for each step.

How does bare metal restore work?

Essentially bare metal restore technology tracks the data that is determined to be necessary to make a particular server operational, such as operating systems, applications, kernels, and registry information. As bare metal restore is initiated, the system is booted and then looks to the source of bootstrap data to rebuild the system based upon the scripted bare metal restore directions. This bootstrap data can be obtained from a network resource, or from a local copy on removable media, depending upon the software's design and capabilities. At the end of the bare metal restore, the system is typically rebooted, and then finds its recovery data from a networked backup server and is automatically restored to a production state. Some backup-software product suites include bare metal restore modules, while others can be augmented with third party utilities that provide this capability.

When to use bare metal restore technology

Bare metal restore technologies are designed to allow for fast recovery of failed systems with less administrator effort, making this technology desirable in any enterprise with RTO pressure. The technology reduces the strain on system administration resources, and can offer improved consistency for system restorations.

Hierarchical storage management (HSM)

HSM and data migration

HSM is software technology that manages the migration of data from a primary storage subsystem to a secondary storage subsystem, based upon defined thresholds. This technology gained popularity in the late 80s and early 90s due to the relatively high cost per megabyte of disk storage as compared to the economical price points of removable storage media, such as optical disk, CD, and tape. As the per megabyte price narrows between optical and magnetic disks, so has the value proposition for HSM. Today, HSM is predominately positioned as augmentation to backup and recovery operations, with the remaining benefit of lower storage cost.

How does HSM technology work?

Traditional HSM technology is positioned primarily as a solution to rejuvenate file systems that have become bogged down with excessive file content, as well as an enhancement to the backup and recovery infrastructure. While backup systems tend to become bogged down with repeated backups of aging, non-active files, HSM provides a means to remove the inactive files from backup requirements. Essentially, as files age in an HSM managed directory, they are migrated onto less expensive storage media. Common second tier HSM storage media options include:

- Optical disk
- CD
- DVD
- Tape
- Low-cost disk drives

Second generation HSM technologies are currently emerging that are integrated with email systems and database management systems to address the management complexities of these rapidly growing data types. By migrating data from those sources, system performance can be improved and backups can be performed more efficiently.

When to use HSM technology

HSM technology can offer relief in backup environments that have become overly burdened with ongoing backup of static data, by moving the files to second tier storage media. HSM technologies generally include their own utilities for backup of migrated data, but this operation is done only once and is less computer-resource intensive, relative to traditional file-based backup systems.

Workstation backup

The need and challenge to backup workstations

Recent research by IDC indicates that up to 60% of the intellectual property owned by an organization is stored on user workstations and laptops, existing outside the protection and management of data center operations. Given that only 18 percent of organizations believe that they have adequate protection for workstation-resident data, it becomes evident that there is a gaping hole in the data protection infrastructure of most organizations. There are a number of reasons behind this wide-scale deficiency:

1. In order to back up workstations using a traditional backup product, the workstations need to be running in a stable state during the backup period, requiring both a disciplined user and a robust workstation operating system.
2. Regular full backups as part of a typical backup procedure require substantial bandwidth, which may be challenging in large organizations with limited backup windows, and relatively impossible for remote users with dial-up connections.
3. Systems rebuild time is prohibitively long for workstations, even with a traditional backup process in place and given system installation and configuration times.

In most organizations, it has been determined that it was too resource intensive to include workstations in the enterprise backup operations. As a workaround, users were generally instructed to copy their backup-worthy files to a network drive, where they would be backed up regularly. In theory, this sounds acceptable, but in practice most organizations find that users lack the discipline to adhere strictly to this standard, leaving countless valuable files unprotected.

Typical workstation backup technology

To address these challenges in recent years, several vendors have introduced technologies that offer improved capabilities for backing up workstations and laptops belonging to the organization. These technologies generally provide the ability to back up workstations and laptops by examining storage blocks on the local hard drive(s) to determine which data has changed, and then only backing up the changed blocks. The backups can occur at any time, and staggering the schedule of backups to occur throughout the production day can minimize bandwidth utilization. The client portions of code generally require very little user interaction or responsibility, which ensures higher backup frequency and reliability given the centralized control.

In addition, the technologies are generally designed to provide block-level incremental backups—where the network utilization due to the backup traffic is minimal. Some products even allow laptop computers to be backed up over a dial-up connection, with the data stream controlled as a percentage of the overall connection capacity to allow continued remote network access, during backup operations.

Continued on next page

Workstation backup, *continued*

New workstation backup technologies

The new workstation backup technologies generally back data up to disk, rather than removable media such as tape. By maintaining the backup workstation and laptop data on disk rather than tape, administrators are freed from the responsibility to load tapes for users who need to perform file recovery. The workstation backup server can be integrated with a network backup scheme to ensure that the backup server is protected in the event of a local disk-drive failure.

For cases of corruption or inadvertent file deletion, users can be granted the ability to perform their own file restorations via a small application on their workstations and laptops. The application can be integrated into existing operating system tools such as Microsoft Explorer. Some products also allow for several versions of a file to be maintained on the server, offering users the ability to restore a previous version of a file.

In the event that a workstation's disk drive requires a complete rebuild, administrator involvement can be greatly reduced when compared to traditional backup methods. Some workstation backup products enable bare metal restoration of computers, using a combination of bootable media with a baseline system configuration, and the ability to transfer and reconcile the remaining files, using the backup server. This could reduce the recovery time by several hours, while the frequent block-level incremental backups ensure that the point of recovery is as recent as possible. See Figure 10.

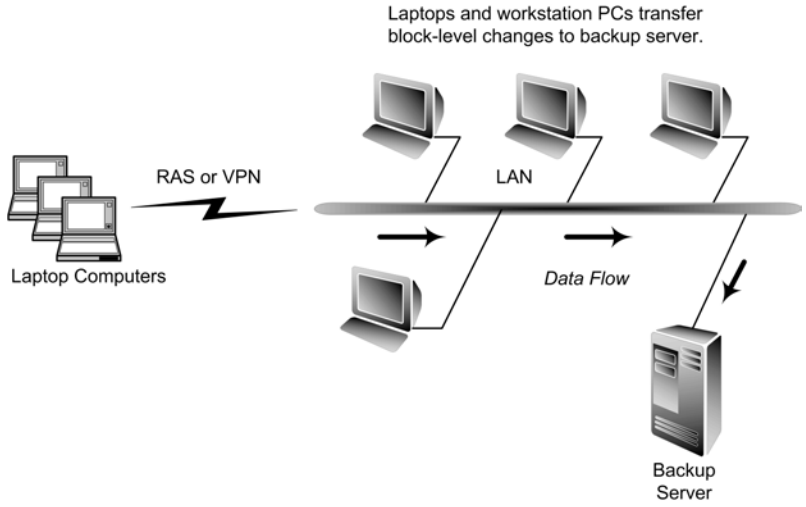


Figure 10: Workstation Backup

When to use workstation backup technology

Any organization with unprotected data on workstations and laptops would benefit from this technology. In addition, organizations that currently use traditional server backup technologies for workstation-resident data can improve their protection coverage and reduce the administrative responsibilities for backup and restoration of this data.

Conclusion

CIOs must become technically informed decision makers as it concerns protecting the organization's data

More than ever CIOs must become more knowledgeable about the discipline of data protection, their storage infrastructures' fault tolerance capabilities, and the technologies that support that end. It is no longer acceptable for CIOs to abstain from intelligent technical discussions with IT staffs about methods of storing and protecting the organizations' most critical business asset, its data.

In today's competitive markets, business data is expanding at rates that outpace the technologies used to manage this growth. That puts tremendous pressure on IT staffs to develop timely, effective, and cost-efficient solutions to meet the storage and protection needs for their evolving and expanding data. Given the need for timeliness and the cost factors associated with storing and protecting data in an information-rich organization, it may be necessary to consider a partnership with an independent storage architect such as Datalink Corporation, to more resourcefully and cost effectively design, implement, and support a customized storage infrastructure. Datalink's expertise is delivering custom storage infrastructures that meet business needs and maximize the value of information.

This paper presented business-driver questions and answers, and technology explanations to prepare CIOs for technical discussions with internal IT staff and external storage architects, such as Datalink, about protecting their organizations' data assets. It is essential that CIOs in today's organizations have this knowledge so that they become technologically informed decision makers as it concerns the process and technologies used to protect their organizations' data assets—to do otherwise, in today's business climate, is not an option.

Terms & Definitions

Terminology

About the glossary The following terms are used in the context of data backup and protection. This glossary serves as a useful resource for processing the information contained within this white paper.

Term	Definition
Backup	A secondary copy of data that is generally used for restoration in the event of damage to the primary copy of data.
Backup Data	The resulting data from a backup operation.
Backup Operation	The process of preparing and copying selected data from primary storage to secondary storage.
Backup Software Components-Client	This portion of software runs on a system that manages a data source to be backed up to a storage subsystem, which is managed by either a primary backup server or a media management server. This code processes the data, adding file-system security information, and prepares the data for transfer to backup media.
Backup Software Components – Media Management Server	This software component resides on a server that maintains a backup storage device, and has read/write privileges on that device. It receives data streams from backup clients and writes the data to the configured backup medium. Metadata regarding each backup session can either be maintained locally on the media management server, or alternatively sent across the network to the primary backup server.
Backup Software Components – Primary Backup Server	This software component serves as the central repository for all metadata pertaining to each backup session. It serves as a master list of all backup history for all files managed within the system.
Backup Types	• Differential backups include all data that has changed within a designated volume since the last full backup. • Incremental backups include all data that has changed within a designated volume since the last incremental backup. • Full backups include all data within a designated volume.
Multiplexing	The process of simultaneously writing multiple data streams to a single tape device; or the process of writing one data stream over multiple tape devices for the purpose of performance enhancements.
Multi-streaming	The process of routing multiple sets of data to a single backup server simultaneously for achieving performance enhancements.
Recovery Point	The point in time at which data is restored. For example, if a backup is performed at midnight, and that data is used to restore a system at noon the next day, the recovery point is midnight. This is a measurement of how much data is lost between the last backup and the time of data loss or corruption.

Continued on next page

Terminology, *continued*

Term	Definition
Recovery Time	The time that it takes to restore a system using backup data, when primary data has been lost or becomes corrupt.
Replication	The copying of data between two services of the same type. Examples are data-to-data service replication or tape-to-tape service replication.
Secondary Storage System	A storage system used for archiving or data protection. Examples are application servers with direct-attached tape drives, libraries, or robots, or dedicated network-attached archiving/data protection appliances.
Storage Units	1 Petabyte = 1024 Terabytes, 1 Terabyte = 1024 Gigabytes, 1 Gigabyte = 1024 Megabytes, 1 Megabyte = 1024 Kilobytes, 1 Kilobyte = 1024 bytes, 1 byte = 8 bits. A bit is the most granular information measurement unit with only two possible values: "0" or "1."
Striping	The process of writing portions of a data set across multiple storage devices to improve data read performance and to enable redundancy.

Datalink Corporation
8170 Upland Circle
Chanhassen, MN 55317
Phone: 1.800.448.6314
www.datalink.com