

**Storage Solutions for Microsoft Exchange Server 2000:
Snapshot Backup and Recovery
with the
IBM TotalStorage™ Enterprise Storage Server**

Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

DISCLAIMERS

The information in this publication is provided “AS IS” WITHOUT WARRANTY OF ANY KIND, implied or otherwise, including but not limited to any warranties of merchantability or fitness for a particular purpose. The responsibility for use of this information or implementation of any of these techniques lies with the customer and depends on the customer's ability to evaluate and integrate these techniques into their operating environment.

Accordingly, the recipient assumes all risks arising from the use of this information. In no event shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity. While IBM may have reviewed each item for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere.

The performance data contained in this document were obtained under a controlled, isolated environment. The results obtained in your operating environment may vary significantly. Accordingly, this performance data does not constitute a performance guarantee or warranty.

Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Copyrights

The copyright of this manuscript is owned by the International Business Machines Corporation. No part of this paper may be reproduced or transmitted in any form without permission.

The following terms are either registered trademarks or trademarks of the IBM Corporation in the United States or other countries or both:

IBM, Enterprise Storage Server, ESS, FlashCopy, Peer-to-Peer Remote Copy, PPRC, Seascape, ESCON, StorWatch, Netfinity, Tivoli, Tivoli Management Environment, TME

The following terms are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries:

Microsoft, Windows, Windows NT, Active Directory

Pentium, Pentium III Xeon, and Intel are registered trademarks of the Intel Corporation.

Other company and product names mentioned herein may be trademarks of their respective companies.



Table of Contents

Executive Summary	4
1 Introduction	5
2 Customer Requirements	5
3 Enterprise Storage Server (codenamed “Shark”) Overview	5
3A Enterprise Storage Server Architecture	6
Figure 1: The Enterprise Storage Server's Internal Architecture	6
3B Enterprise Storage Server Advanced Copy Services	7
4 Microsoft Exchange Server 2000 Features and Requirements ..	8
4A Exchange Server Backup and Recovery Overview	8
4B Exchange Server Components and Considerations	9
4C Using Exchange Server with a Secondary Server	11
5 Snapshot Backup and Recovery Setup and Configuration	13
5A Storage Configuration	13
5B Hardware Configuration	13
5C Network Topology	14
Figure 2: The Storage Area Network Physical Topology	14
5D Snapshot Infrastructure Setup	15
Figure 3: The Storage Area Network Logical Topology	15
6 Snapshot Backup Process	16
6A Backup Scenario 1 - Local FlashCopy Backup	17
Figure 4: Backup Scenario 1 - Local FlashCopy Backup	17
6B Backup Scenario 2 - Remote FlashCopy Backup	18
Figure 5: Backup Scenario 2 - Remote FlashCopy Backup	18
7 Recovery	19
7A Recovery Scenario 1 - PPRC Recovery	20
Figure 6: Recovery Scenario 1 - PPRC Recovery	20
7B Recovery Scenario 2 - FlashCopy Point-in-time Recovery	21
Figure 7: Recovery Scenarios 2 & 3 - Local FlashCopy Restore	21
7C Recovery Scenario 3 - FlashCopy Roll-Forward Recovery	22
8 Process Automation and Solution Integration	22
Appendix A: Hardware and Software Configuration	23
Appendix B: Operating System Considerations	24
References	26



Executive Summary

As messaging and collaboration become mission critical for more corporations, these corporations need an enterprise-class storage subsystem to support their messaging platform. Microsoft Exchange Server is rapidly becoming the platform of choice for some companies. For Microsoft Exchange customers, the IBM TotalStorage Enterprise Storage Server (codenamed "Shark") can provide the reliability, scalability, and features required for a mission-critical application.

This white paper describes how a customer can use the Enterprise Storage Server's Advanced Copy Services to perform instant snapshot backup and recovery in a Microsoft Exchange Server 2000 environment. One of these Copy Services, FlashCopy, can perform serverless backups of multi-terabyte databases in minutes. It can also help improve uptime by providing the ability to recover in minutes, instead of hours or days. Another Copy Service, Peer-to-Peer Remote Copy, provides a synchronous remote mirror to protect from disasters. When used in combination, these two Copy Services have the ability to provide multiple instant copies of the database (for serverless backup, application development or testing, and other functions) and can be used as a part of an overall plan to protect against disaster.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

1 Introduction

With the release of Exchange Server 2000, Microsoft has unleashed a collection of new functions for their flagship messaging platform. Beyond e-mail, companies rely on Exchange for collaboration and web applications, and increasingly, conferencing and instant-messaging. With all of these features, Exchange has become a mission-critical component for many companies. In short, enterprises depend on Exchange to get their work done.

As with any mission-critical application, Exchange requires an enterprise-class storage management solution for improved availability and performance. This paper describes the use of the IBM TotalStorage Enterprise Storage Server (ESS) and its Advanced Copy Services to perform snapshot backup and recovery for Exchange Server using a set of documented scenarios. These techniques are intended to leverage features of the ESS and Microsoft Exchange Server to allow the near-instantaneous creation of copies of data, without compromising either data integrity or performance of online operations.

2 Customer Requirements

Backup Requirements:

As customers rapidly move to a 24x7 global computing environment, the backup window is shrinking just as rapidly. Customers can no longer afford the daily downtime on their production server to perform backups. This becomes especially true as databases become larger and larger. Online backups improve the availability of the database, but costs valuable host CPU resources, disk resources, and network resources. Even with the use of incremental and differential backups, the time necessary to perform a backup is significant.

Recovery Requirements:

Should an error occur or disaster strike and recovery is needed, a restore often takes twice the amount of time as a backup (or more). During this time, the database will be unavailable. While backup techniques such as incremental and differential backups reduce the backup time significantly, they actually *increase* the amount of time required to restore.

Despite planned and unplanned outages, business needs often require that databases must be available within minutes. While disasters and hardware failures are rare, logical errors and software errors are more frequent. These errors require a time-consuming restore as well.

As databases become larger and contain more mission-critical data, they increasingly require higher availability. Conversely, the time required to back up and restore becomes longer.

3 Enterprise Storage Server (codenamed “Shark”) Overview

To help meet the challenges inherent in an enterprise-class database infrastructure, IBM introduced the Enterprise Storage Server (ESS), codenamed “Shark.” Its rock-solid reliability, high performance, and Advanced Copy Services make it well equipped to meet the challenges.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

The ESS provides enterprise-class storage for the applications across an entire corporation. Based on a third-generation Seascape architecture, the ESS is designed to deliver scalability, availability, performance, and connectivity to nearly every server in the enterprise. Availability and performance can be further improved by using its Advanced Copy Services features, including FlashCopy and Peer-to-Peer Remote Copy (PPRC).

3A Enterprise Storage Server Architecture

IBM designed the ESS specifically for the enterprise solutions market, which requires high reliability and availability. IBM's focus on reliability and availability clearly shows in the advanced architecture and features of the ESS:

Advanced Internal Architecture:

Each ESS consists of two complete storage subsystems attached to fault-tolerant disks in a clustered configuration. Each of the two subsystems, or clusters, contains two buses to connect to the hosts, and four internal storage controllers to attach to the disks. The ESS also contains vast amounts of cache, dedicated nonvolatile storage (NVS), and intelligent load-balancing and parallelism. The result is a high performance, highly available storage server that is designed to avoid single points of failure. Figure 1 shows the multiple paths to data available within the internal architecture of the ESS:

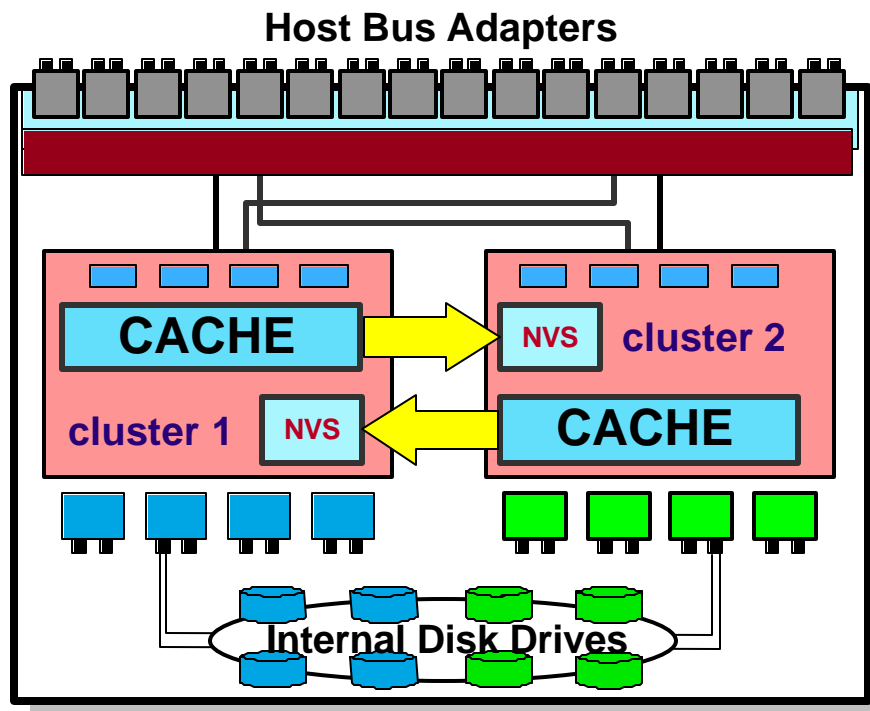


Figure 1: The Enterprise Storage Server's Internal Architecture

Advanced External Connectivity:

The ESS can connect to multiple host systems using multiple Fibre Channel / FC-AL or SCSI connections. These connections can be multipathed for greater performance and redundancy using IBM's Subsystem



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

Device Driver (SDD). The SDD software, installed on the host, automatically balances the I/O load across all available connections from the host to the ESS. In case of a connection loss, such as during a host bus adapter failure, SDD is designed to automatically failover to another path. These features support the host servers' ability to have fast, continuous access to data.

Other ESS Features:

These features, along with non-disruptive upgrades and repairs, help provide the availability required in an enterprise-class solution. For more information about the ESS and other features of the ESS (such as SAN-capability, ease of management, and universal data consolidation), please visit the IBM storage web site at <http://www.storage.ibm.com>.

3B Enterprise Storage Server Advanced Copy Services

To further improve availability and performance, the IBM ESS features Advanced Copy Services. These serverless copy functions are implemented entirely within the storage server, with little or no impact on the host server.

FlashCopy:

FlashCopy is designed to provide an instant T0 copy of data. A T0 copy is a point in time copy of the storage volumes when the FlashCopy was invoked. The copy is available almost immediately after invoking the command. The copy process is transparent to the host server, as the ESS storage subsystems manage the copy internally. The point in time copy can now be accessed by a secondary server and be backed up to tape, typically with no significant impact on the host server.

With FlashCopy, multi-*terabyte* databases often can be backed up and restored in minutes, instead of hours or days. Even for a tape subsystem capable of 100 GB/hr, a 1 TB database might take ten hours to backup. A restore from tape will take at least the same amount of time. Using the FlashCopy, the same database can be backed up almost instantaneously. Furthermore, the database often can be restored in minutes. Aside from backup and restore, FlashCopy can also provide a copy of "near-live" data clone for business intelligence or data mining, application testing and development.

When FlashCopy is invoked, the ESS creates another copy of the data by building a bitmap that records changed data. When the bitmap is complete (typically in seconds), the copy is *logically complete*. Both copies can immediately be used (read and write) separately without affecting the other copy. The ESS then begins physically copying the data to the target set of disks (at a rate much faster than a host server could copy). Whenever the host server writes to a block that has not yet been copied, the data is copied first, and then the write continues. In this way, an identical T0 copy can be made almost instantly with little or no impact to the host server.

Peer-to-Peer Remote Copy:

PPRC is a remote mirroring technology (similar to RAID-1) that is designed to provide a synchronous copy of production data at a remote site. Because the mirroring is done at the storage subsystem level, there is almost no impact on the host. The remote site can be up to 103 km away, or more with channel extenders. In case of a disaster at the production site, such as fire or earthquake, an up-to-the-second copy can be available at the remote site. The remote copy can be brought online using a standby server, with minimal business interruption.



4 Microsoft Exchange Server 2000 Features and Requirements

Exchange Server 2000 contains many features that are radically different from its predecessor, Exchange Server 5.5. While these features offer many functional enhancements, they also complicate the backup and restore process significantly. Some of these changes are discussed below.

4A Exchange Server Backup and Recovery Overview

Exchange Server DBAs today are familiar with the various types of traditional tape and disk based backups. For completeness, and to create a point of reference for terminology, a brief discussion of the various types of backup and recovery ensues. While FlashCopy and PPRC change *how* and *how fast* backup and recovery is performed, the principles and purposes (the *what* and the *why*) are the same.

Backup Types:

Microsoft Exchange Server supports several different types of backup, including differential and incremental (log) backups. However, for the purposes of snapshot backups, only full (non-incremental) backups are supported. In addition to full backups, log backups save the data that has changed since the last full or log backup. We recommend that log backups be taken in between full backups to further reduce exposure to data loss. We also recommend periodically truncating the logs (typically after the full backup) to prevent the logs from growing without bound.

Currently, Exchange Server 2000 does not offer a way of ensuring consistency during an online snapshot backup. Such a feature may be offered in a future Service Pack from Microsoft. In the meantime, it is necessary to shut down the database briefly to perform a snapshot backup. Customers who require 24x7 availability of their messaging platform can use FlashCopy to restore in minutes instead of hours.

In the absence of true online snapshot backup capability in Exchange 2000, we recommend that customers use Copy Services in conjunction with traditional log backups to help improve availability and performance of their Exchange 2000 servers. As long as the customer can tolerate a brief outage to perform nightly or weekly backups, recovery time can be greatly reduced.

Recovery Types:

Exchange Server supports two types of recovery: point-in-time and full roll-forward. Point-in-time recovery is a restore up to the time of the last backup. All transactions since the last backup are lost. This type of recovery is useful for recovering from software errors or human errors. It can employ circular logging to minimize the amount of space used for transaction logs. Roll-forward recovery is a restore up to the last transaction available (usually the point of failure). As long as the transaction logs are not damaged, no data should be lost.

Designing and implementing a backup and restore strategy is lengthy, complex task, and is outside the scope of this document. For an introduction to designing a backup and restore strategy, please see the IBM Redbook *Using Tivoli Data Protection for Microsoft Exchange Server* available at <http://www.redbooks.ibm.com>.



4B Exchange Server Components and Considerations

Storage Groups and Databases

Exchange data is stored within databases, which in turn is stored in storage groups. Exchange 5.5 featured one storage group with two databases, the public (pub.edb) and private (priv.edb) information stores. To improve scalability, Exchange Server 2000 Enterprise Edition features up to four storage groups, each of which contains up to five databases for a maximum of twenty databases. Each storage group is a separate process running store.exe. Each storage group also contains one set of transaction logs that are shared by all of its databases. Each database can be backed up and restored independently of all the others in the storage group. For example, in an organization with 2000 users spread evenly between twenty databases, if one database is corrupt, it can be taken offline and restored while the other 1900 users continue seamlessly.

With this capability, you can divide your organization into “time zones” and perform backups and restores without impacting other “time zones.” For example, say you have an office in Asia and an office in the U.S. Since they are about 12 hours apart, you could take down the storage group in Asia and back it up at midnight local time while the US database continues without interruption. Databases are usually mounted and dismounted from the Exchange System Manager GUI. In the lab, we developed a command-line utility to perform these actions, so that backup tasks can be easily automated.

In reality, backups are much easier to manage on a storage group level. While each database can be backed up and restored independently, only one backup or restore can occur at a time per storage group. Furthermore, transaction logs are maintained on a storage group level, and it is difficult if not impossible to extract the portion of the logs pertaining to a particular database to back up. Therefore, while restore works well at the database level, we recommend performing backups on a storage group level.

Each Exchange Server 2000 database consists of two database files, the EDB database file and the streaming STM database file. In Exchange 5.5, all Internet messages (MIME format) were converted to MAPI format and stored in the EDB file. This conversion may have an adverse affect on performance. With Exchange Server 2000, Internet messages are stored in the STM database file, in its raw or streaming format. The conversion is only required for MAPI clients. The EDB file then contains a pointer to the location inside the STM file. While this greatly enhances performance, it adds another dimension to backup and restore: the EDB and STM files are inextricably linked, and both files are necessary for a restore. If either file is missing, the database will not start. Microsoft provides the utility, "eseutil.exe", to determine whether the EDB file and the STM file are a matched pair.

Transaction Logs

As with any modern database, Exchange uses transaction logs for additional performance and protection against failure. As soon as data is received by Exchange, the information is written to the transaction logs. The data will then be asynchronously written to the datafiles as load permits. Until the transaction is destaged to the datafiles, access is done in memory. Once the log data is destaged, that particular log is no longer used except for recovery. It is very important to backup the transaction logs, as they are required to roll forward after a recovery, or to recover from a backup prior to the last backup.

Exchange logs are uniformly 5 MB in size. The logs are named sequentially, in hex, starting at Exx00000.log. The Exx indicates the storage group instance, typically E00 to E03. The current log being used by Exchange is called Exx.log. When that log fills up, it is renamed to the next sequential log, and



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

Exchange begins writing to a new log, named Exx.log. Therefore, at any time, Exx.log is the only log being written. The number of logs will continue to grow until truncated. It is important that the logs be periodically truncated (but only after being backed up safely!). If the logs are not maintained properly, they will continue to grow until there is no more space, and Exchange shuts down. Exchange keeps two reserve logs (res1.log and res2.log) in case it runs out of space on the drive. In that case, Exchange will write out the remaining transactions to the reserve logs and shut down immediately.

Typically, the transaction logs are truncated at backup time. Only the logs that have been destaged to the data files can be truncated. The other logs should not be truncated, as Exchange still needs to commit the data to the datafiles. In the lab, we developed a program to automatically determine which logs had been destaged, back them up using FlashCopy, verify the backup integrity, and then truncate them. These steps are described in detail in Section 6, "Snapshot Backup Process." Truncating the log without first backing them up is risky. Should the production database fail, and the backup becomes corrupted, it is only possible to restore up to the time of the last good backup. Any changes since then would be lost.

One way of saving disk space is to use circular logging. Instead of archiving the log, Exchange can overwrite logs that have already been destaged to the datafiles. However, without the logs, the database is recoverable only up to the time of the last good backup. Any transactions that have occurred since then will be lost. Circular logging is generally not appropriate for mission-critical data. Therefore, by default, circular logging is disabled.

Checkpoint File (Exx.chk)

The checkpoint file, stored in the system directory, indicates which log file was last destaged into the data file. Each storage group has its own checkpoint file. To perform a point-in-time recovery (to recover from a user error, for example), be sure to back up the checkpoint file. Without the checkpoint file, the recovery process must perform a time consuming traversal of all available transaction logs, and it will continue until the last available log (perhaps past the desired point-in-time).

Active Directory

Active Directory is Windows 2000's database for maintaining user and computer information, and other information. In Exchange 5.5, the directory store maintained the list of Exchange users and their mailboxes. With Exchange Server 2000, the directory store is replaced by the Active Directory. Just as the directory store was an integral and essential part of Exchange 5.5, the Active Directory is an integral and essential part of Exchange Server 2000. Without an Active Directory, Exchange will not run. Accordingly, the Active Directory must be backed up separately and regularly.

Internet Information Server Integration

The previous version of Exchange used X.400 as its native transport protocol. The new version of Exchange, Exchange Server 2000, uses Internet-standard SMTP and POP3. These services are provided by the Internet Information Server (IIS), a component of Microsoft Windows NT and 2000. Therefore, IIS is a prerequisite to installing Exchange Server 2000.

IIS contains a metabase that is necessary for Exchange Server 2000. However, since most of the data is replicated inside Active Directory, the IIS metabase often does not need to be backed up. In cases where IIS is heavily tuned or configured, or contains complex NNTP information, it should be backed up as part of a system state backup.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

Key Management Server Database

The Key Management Server (KMS) database is the recovery database for encryption keys. Users receive an encryption key when they send an encrypted e-mail. Should the user lose the key, the KMS database can be used to recover the key. If the user loses the key and the KMS database is not available, then the data is permanently lost (not decipherable). For this reason, it is important to be able to recover the KMS database. The use of encryption keys and KMS is outside the scope of this document.

Site Replication Service Database

The Site Replication Service (SRS) database causes an Exchange Server 2000 Server to appear as an Exchange 5.5 Server, so that the two can coexist. If possible, back up the SRS database. If the SRS database is not available at the time of restore, it can usually be recreated, but this can consume a lot of resources as data is replicated across servers. The use of the SRS database is also outside the scope of this document.

4C Using Exchange Server with a Secondary Server

Aside from backup and recovery, a customer may wish to use ESS's Copy Services to copy an Exchange Server 2000 database to a secondary server and connect users to that secondary server. There are several issues to consider to ensure a successful copy.

Active Directory Considerations

As mentioned earlier, Exchange Server 2000 is not usable without an Active Directory. In previous versions of Exchange, it was possible to copy both the Directory store and the Information store to a secondary server. With Exchange Server 2000, each Active Directory is unique, so it cannot simply be copied to the secondary server (except perhaps if the secondary server is not connected to any network). As a workaround, Microsoft provides a utility, mbconn.exe, that makes it possible to use the database on the secondary. Before using the utility, install Exchange Server 2000 in a separate Active Directory forest. The server can remain connected to the network because the users are created in a different forest.

Mailbox Reconnect Tool (mbconn.exe)

Microsoft provides the utility mbconn.exe on the Exchange Server 2000 CD-ROM in the \Support\Utils\i386 folder. It uses a two step process. First, it generates Active Directory users on the secondary machine. The utility gathers information about users from the Exchange database and stores it in a .ldf file. The Active Directory Import/Export tool (ldifde.exe) then imports that file into the new Active Directory. Secondly, after the users are all imported into the new Active Directory (typically in seconds), mbconn.exe connects those users to mailboxes in the Exchange database. Once the reconnection is completed, users can log on and use the copied mailboxes. For more information about the mbconn.exe utility, please see Microsoft's knowledge base article Q271886.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

Legacy Exchange Distinguished Names (legacyExchangeDN)

Another issue to consider is the naming of the Exchange server. To successfully use the database on a secondary server, be sure that the legacyExchangeDN stem values matches that of the production server. In general, this can be done by matching the following names when installing the secondary server:

- a. Exchange Server 2000 organization name
- b. Administrative Group name
- c. Storage Group Name
- d. Database name

To view the legacyExchangeDN name, use the Active Directory Import/Export tool (ldifde.exe) to export a few users to a file. The legacyExchangeDN has the following form:

legacyExchangeDN:

/O=organization_name/OU=administrative_group_name

It is possible to change the legacyExchangeDN, but the process is labor intensive. It is much easier to install Exchange using the correct names.

Single Mailbox Recovery

In some cases, an Exchange administrator may wish to restore a single mailbox instead of an entire database. For example, if a user accidentally deletes some important messages, an administrator might want to restore the mailbox to a previous point in time (before the deletion). If the entire database is restored, though, other users will be reverted to a previous point in time. As this is unacceptable, Exchange provides a few mechanisms to help alleviate the problem. Firstly, Exchange retains deleted messages for a configurable amount of time, such as for seven days. During this time, the user can recover his or her own messages. After the retention period, Exchange deletes the messages.

At this point, the data is recoverable by using a second recovery Exchange server and a previous backup. Perform the following steps on the recovery Exchange server:

1. Set up the recovery server using methods/considerations described above
2. Restore previous copy of Exchange (using Copy Services or previous backup)
3. Run the Mailbox Cleanup Agent (right-click on Mailboxes, choose Run Cleanup Agent)
4. Create a non-mailbox enabled Active Directory user account
5. Reconnect mailbox to new Active Directory user (Right click the user's mailbox, select Reconnect)

After these steps, the messages are available on the secondary Exchange server. There are several ways to get the messages back on to the original Exchange server. Probably the easiest way to do this is to use Exmerge.exe 2000 (available on the Exchange 2000 CD) to export the messages from the secondary Exchange server. The messages are saved in a .pst file. Then, use Exmerge.exe on the original Exchange server to import the messages to the appropriate mailbox.

For more information about single mailbox recovery, legacyExchangeDNs, and other recovery issues, please refer to Microsoft's white paper, "Microsoft Exchange Server 2000 Server Database Recovery."



5 Snapshot Backup and Recovery Setup and Configuration

To validate these methods of snapshot backup and recovery, we applied a set of test scenarios in the lab. We created an environment that was intended to closely resemble a typical customer environment, while at the same time heeding Microsoft's "Best Practices" recommendations, published in previous white papers.

5A Storage Configuration

In general, it is important to place the transaction logs and the data files on separate physical disks or volumes. Though the ESS is designed to be fully fault-tolerant, this separation provides yet another layer of protection. Should the data files be lost, it would still be possible to recover up to the point of failure using the log files and a backup of the data files. Also, there can be performance benefits in placing these files on separate media. With this in mind, we set up the Exchange database to spread the data and logs across all of the available disks in the ESS for best performance. We also spread the load across both clusters, all host bus adapters, and all internal logical subsystems (LSSs) available within the ESS. This helps avoid "hot spots" within the storage subsystem that receive most of the load while other parts of the subsystem are idle.

Furthermore, all of the transaction logs were mirrored on a separate set of disks. The ESS is designed to avoid single failures (such as a disk, adapter, CPU, or memory module failure), and most double and triple failures, with its redundant, highly available architecture. The possibility exists, however remote, that a combination of hardware failures can occur that would cause data loss (usually during a disaster). Having a mirror of the log is extra protection against such a scenario. Additionally, in some scenarios, we have a mirror at a remote site to protect against disasters, such as an earthquake or fire.

5B Hardware Configuration

The test environment contained two Intel-based IBM Netfinity Servers and two Enterprise Storage Servers all connected through an enterprise-class McData ED-5000 Fibre Channel Director. Using Fibre Channel, the Storage Area Network (SAN) provided a fast path to data without adding congestion to the IP network. Each of the Netfinity Servers connected to the SAN via two Fibre Channel host bus adapters (HBAs). Each of the ESSs connected to the SAN using four Fibre Channel adapters, and to each other over 8 ESCON links. In total, there were eight redundant logical paths from each host to each ESS. Each server was configured to use four of the eight available paths. The paths were logically combined through the use of IBM's multipathing software, SDD. They appeared to Windows as a single logical path (i.e., one instance of each volume). The server configurations are shown in Appendix A.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

5C Network Topology

These servers were connected into a Storage Area Network (SAN). The SAN had the following topology, shown in Figure 2:

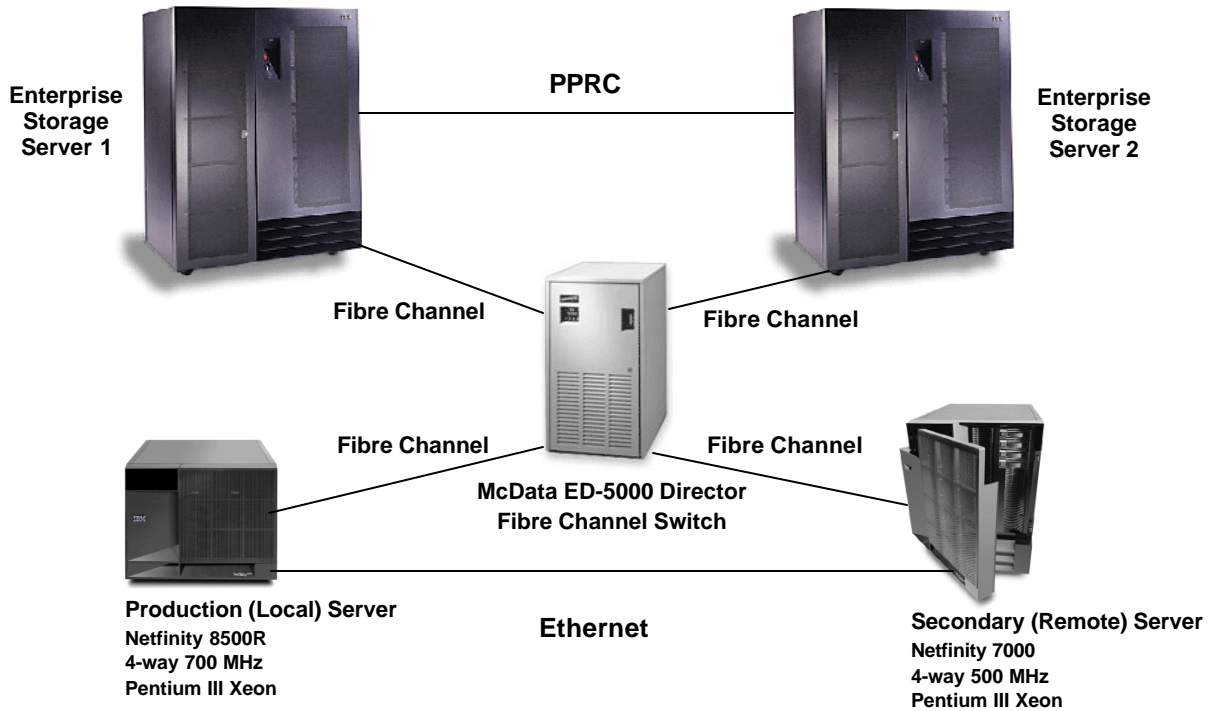


Figure 2: The Storage Area Network Physical Topology

In this configuration, there is the potential for a single point of failure at the McData switch. Any fibre channel switch, even an enterprise-class, fully redundant Director, can represent a single point of failure if there is a disaster or multiple failures. The McData switch is meant to represent a SAN fabric, consisting of several redundant paths and switches from edge to edge. In that way, single points of failure are avoided.

Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

5D Snapshot Infrastructure Setup

To attain higher availability, we devised a two-ESS and two-server configuration. The production server and ESS 1 are located at the production site, of course. The secondary server and ESS 2 are located at a remote site. Figure 3 is a logical view of the environment configuration:

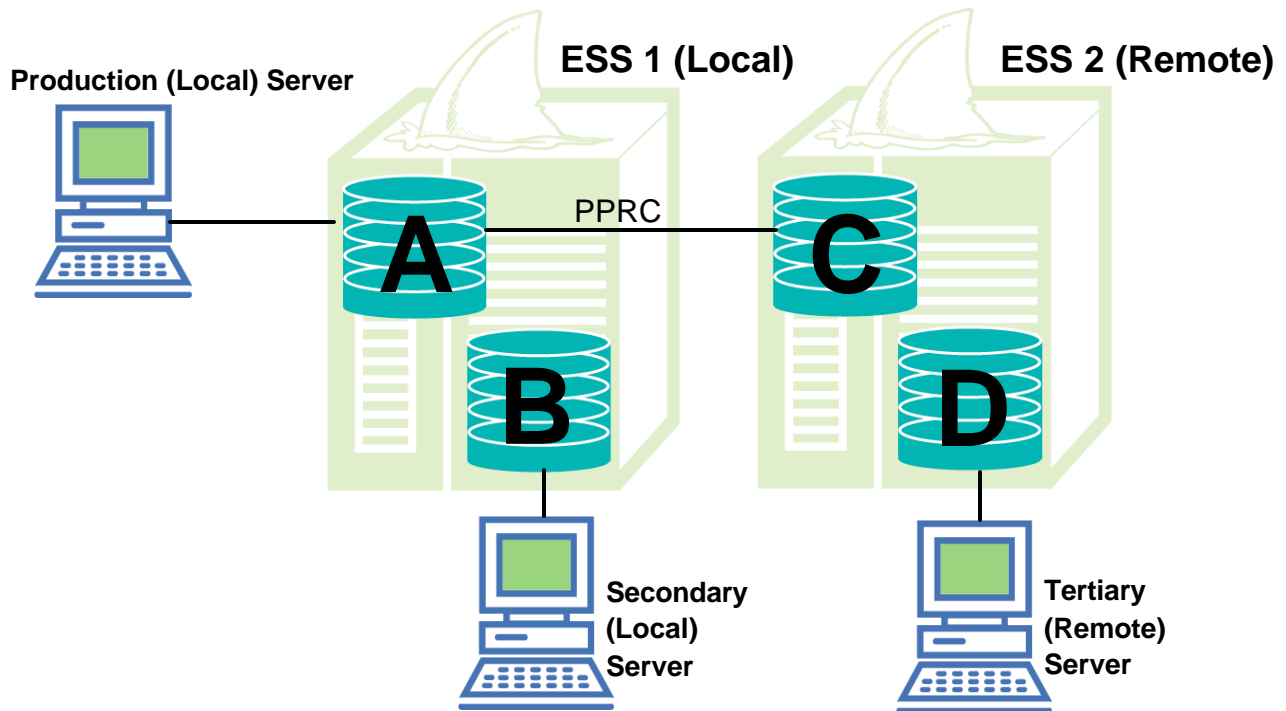


Figure 3: The Storage Area Network Logical Topology

At the production site, the ESS contains two copies of the production database. The first copy (A) is the production database. Periodically (perhaps every night), the database is quiesced and FlashCopy (B) is created. The FlashCopy (B) is then sent to tape using the resources of a secondary server (not the production server). After that, the FlashCopy (B) is maintained in case a rapid restore is required. At the remote site, the second ESS keeps a synchronous copy (C) of the production data. Should the primary server fail, the backup server at the remote site can take over, typically with little or minimal business interruption. It is also possible to take a FlashCopy of that remote copy to produce a fourth copy (D) that can be used for additional protection, data mining or business intelligence, application development, etc.

Customers may choose to have a secondary server at the local site to send the backup to tape, and a tertiary server at the remote site to take over in case of disaster. In our test environment, our second server served as both the secondary and tertiary.

6 Snapshot Backup Process

This section discusses two situations in which customers might use FlashCopy and PPRC to perform backups. Restore scenarios are discussed in Section 7, “Recovery”. Using FlashCopy and PPRC, it is possible to make dozens of copies in different configurations. We devised two likely scenarios a customer might use to perform a database backup.

Scenario 1 : Customer requires serverless fast backup and recovery

Solution : Offline local FlashCopy backup

As with each of these scenarios, Scenario 1 provides an instant backup copy of the database. For customers who require the ability to recover as quickly as possible, this scenario allows for a fast FlashCopy restore.

Scenario 2 : Customer requires disaster recovery + Scenario 1 requirements

Solution : Offline remote FlashCopy backup

Using PPRC and FlashCopy, this scenario provides an instant copy on a second ESS. This scenario adds further protection and reduces or eliminates any impact on the primary storage subsystem. With the second ESS, the database can typically survive a total site failure.

While these backup scenarios performed as expected under significant load in the lab, we recommend that the backup be performed at times of minimal load (such as during the night). In each case, a standby Exchange Server had been installed at the secondary server. Furthermore, each of our scenarios included error detection mechanisms. We recommend that customers use rigorous error checking in their implementation of these scenarios.

Warning: Do not start Exchange Server on the secondary system if the secondary database will be used for backup and restore. Instead, use an application such as NTBackup or the Tivoli Storage Manager Backup-Archive client to make a flat file-level backup. Restarting Exchange Server can modify the database so that it can no longer be used for roll-forward recovery. The steps for starting the database on the secondary should only be used for application development or testing.

Each of these scenarios are described in further detail in subsequent sections.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

6A Backup Scenario 1 - Local FlashCopy Backup

This scenario requires one ESS at the production site. While this scenario provides very fast backups and restores, it does not offer any protection against disasters, such as fire or earthquake. Logically, the data is being copied from copy (A) to copy (B) in Figure 4 below:

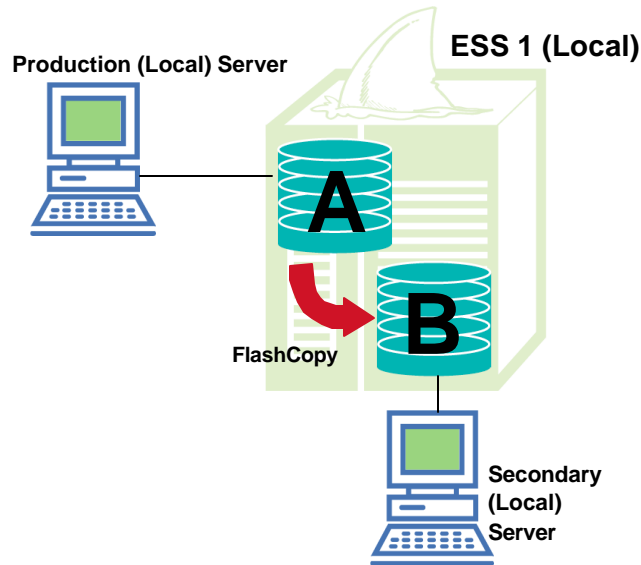


Figure 4: Backup Scenario 1 - Local FlashCopy Backup

Customers that require a known, consistent state, and can tolerate a brief outage should use an offline FlashCopy backup. The steps for performing the backup are as follows:

1. Stop production Exchange services (POP3, IMAP4, etc.), and Information Store
2. Flush file system buffers on production server and secondary server
3. Perform FlashCopy backup
4. Resync secondary filesystem (i.e. make secondary server aware of new disks)
5. Verify log backup (eseutil /ml)
6. Truncate production server's logs
7. Restart production Exchange services (POP3, IMAP4, etc.), and Information Store
8. Check backup database integrity, consistency, and signatures at secondary server
9. (optional) Restart secondary Exchange services (POP3, IMAP4, etc.), and Information Store

In the lab, we developed a program that, after safely performing the backup, automatically truncates the logs that have already been destaged to the data files. Scripts performed the remaining steps and error checking.



6B Backup Scenario 2 - Remote FlashCopy Backup

To add protection from disasters (such as fire, earthquake, etc.), this scenario requires one ESS at the production site and one ESS at the remote site for disaster recovery. The FlashCopy is done from the remote synchronous mirror (C) to the remote copy (D), as shown in Figure 5:

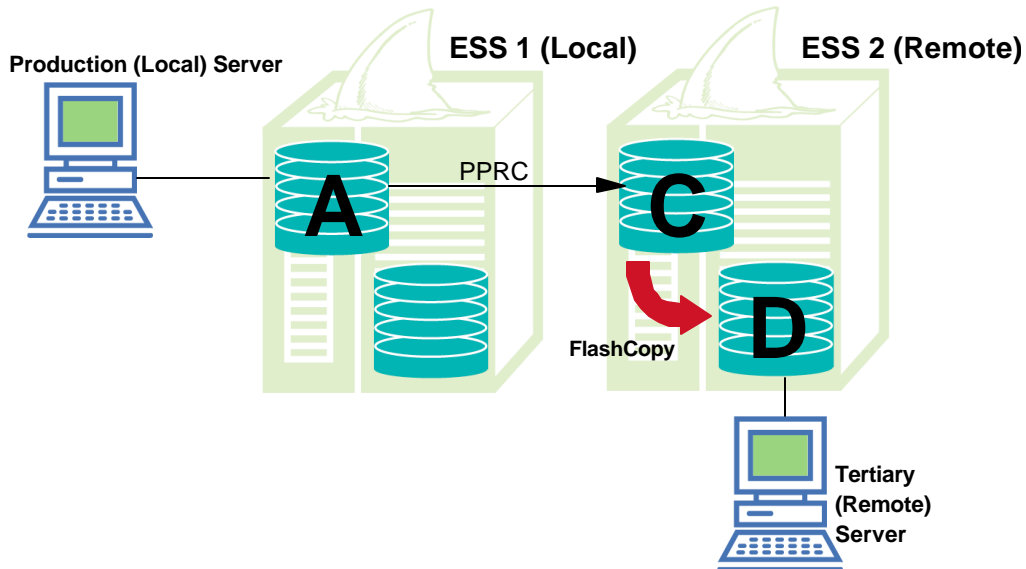


Figure 5: Backup Scenario 2 - Remote FlashCopy Backup

This setup provides the rapid backups functionality of Scenario 1 and adds disaster recovery protection. Furthermore, the FlashCopy (D) of the database at the remote ESS (ESS 2) can be used to send the backup to tape, eliminating any overhead on the production ESS and production server. Before running the scenario, the PPRC source and target volumes must be in duplex mode (i.e. full volume copy is completed and writes are synchronized). The following steps are involved in the offline remote FlashCopy backup:

1. Stop production Exchange services (POP3, IMAP4, etc.), and Information Store
2. Flush file system buffers on production server and remote server
3. Perform FlashCopy backup at remote ESS 2
4. Resync remote filesystem (i.e. make remote server aware of new disks)
5. Verify log backup (eseutil /ml)
6. Truncate production logs
7. Restart production Exchange services (POP3, IMAP4, etc.), and Information Store
8. Check backup database integrity, consistency, and signatures at remote server
9. (optional) Restart secondary Exchange services (POP3, IMAP4, etc.), and Information Store

A customer can combine this scenario with Scenario 1 for both disaster protection and extremely rapid restores. This combination further improves availability.

7 Recovery

In addition to the two backup scenarios, we devised three recovery scenarios that can be used to recover from most errors or disasters.

Scenario 1 : Customer requires disaster recovery

Solution : PPRC recovery

In the situation where the production database is corrupted or the production ESS is lost (due to natural disasters, etc.) we can use the backup taken in Backup Scenario 2 to restore from the remote copy on ESS 2. While the synchronous nature of PPRC protects against disasters, it does not protect against logical errors; the errors are propagated to the PPRC Target on the remote ESS as well. This scenario is designed to protect against logical errors. It can also help protect against rolling disasters. The FlashCopy taken at the remote site is a known consistent backup, and can be used to restore the production database much faster than possible from tape. Optionally, if the log files are not damaged and contain no logical errors, they can be used to roll forward to the point of failure.

Scenario 2 : Customer requires point-in-time recovery

Solution : FlashCopy point-in-time recovery

In the scenario where the production database is corrupted or contains logical errors, the customer would need to restore the database to a previous point-in-time from the FlashCopy backup. The FlashCopy backup taken in Scenario 1 contains a known, consistent state, and is used to restore the database.

Scenario 3 : Customer requires recovery to point-of-failure

Solution : FlashCopy roll-forward recovery

This scenario would be used by customers who can not afford to lose any data in case of a system failure. As long as the transaction logs are not damaged, this scenario can be used to recover the database up to the point of failure. This scenario, combined with frequent log backups, offers a very short restore window with no data loss.

Before attempting any recovery, we recommend making a copy of the transaction logs. If time permits, we also recommend backing up the data files. Even if the data files are not startable, they may be repairable. While each of these scenarios have been fully validated in the lab, the backup provides yet another safety net for mission-critical data.

In each scenario, it is assumed that Exchange Server is already installed and all pertinent patches applied. Before executing any restore, be sure to select the **This database can be overwritten by a restore** check box in Exchange System Manager, in the Database properties of the database object.

Each of these scenarios are described in further detail in subsequent sections.



7A Recovery Scenario 1 - PPRC Recovery

This scenario requires two ESSs and assumes the database setup of Backup Scenario 2. Backup Scenario 2 makes periodic (perhaps daily) FlashCopy backups of the PPRC target on the remote ESS, and this backup is used to restore the production database on the primary ESS. Figure 6 shows the logical view of the scenario:

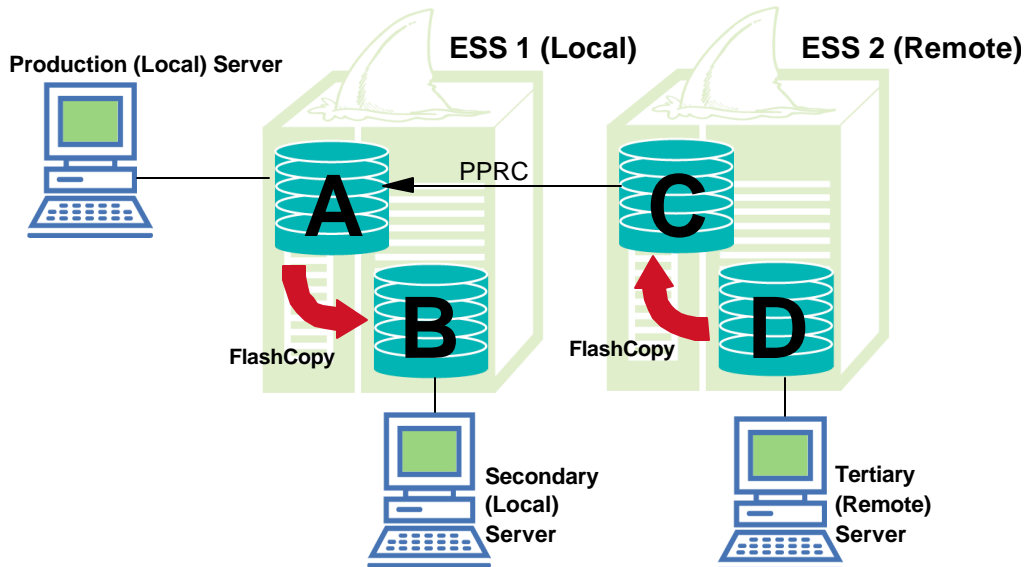


Figure 6: Recovery Scenario 1 - PPRC Recovery

This scenario assumes that there is no local copy available. The steps in detail are:

1. Stop production Exchange services (POP3, IMAP4, etc.), and Information Store
2. Terminate PPRC relationship from production ESS 1 to remote ESS 2
3. Flush file system buffers on production system and remote system
4. Perform reverse FlashCopy at remote ESS 2 from copy (D) to copy (C)
5. Establish PPRC relationship from remote ESS 2 copy (C) to production ESS 1 copy (A)
6. Wait for PPRC to complete (check status with rsQueryComplete)
7. Terminate PPRC relationship from remote ESS 2 copy (C) to production ESS 2 copy (A)
8. Resync production filesystem (i.e. make production system aware of new disks)
9. (optional) Check backup database integrity, consistency, and signatures at production server
10. Start production Exchange services (POP3, IMAP4, etc.), and Information Store
11. Perform a safety FlashCopy at production ESS 1 from copy (A) to copy (B)
12. Re-establish PPRC relationship from production ESS 1 copy (A) to remote ESS 2 copy (C)

Once the database is restored successfully (after step 10), we take a safety FlashCopy (B) to be used for rapid restores in case the production database becomes corrupt again. Then, at the conclusion of this scenario, we restore the ESS configurations to the production state, with the production database (A) being mirrored on the remote site (C) through PPRC.

7B Recovery Scenario 2 - FlashCopy Point-in-time Recovery

While Recovery Scenario 1 provides protection against disasters, its method of restore can take several hours for very large databases. Though PPRC is usually faster than multiple tape drives, FlashCopy can provide even faster restores. With this scenario, the customer is able to restore to the time of the last backup within a very short period of time. In this scenario, as well as Recovery Scenario 3 (roll-forward recovery), the logical setup is as shown in Figure 7:

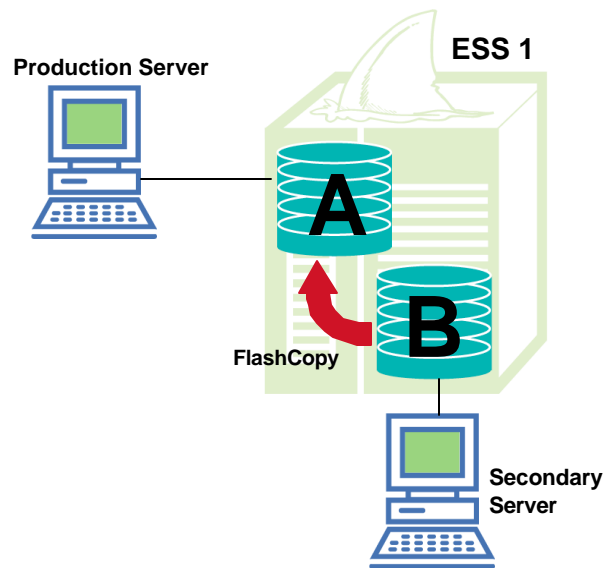


Figure 7: Recovery Scenarios 2 & 3 - Local FlashCopy Restore

This scenario is useful for recovering from logical errors. If an administrator accidentally deletes or corrupts a database, for example, a FlashCopy can return the database to a consistent, known prior state (the state at the time of the last backup). Any changes made after that last backup need to be redone (except the ones that caused logical errors, of course).

Exchange Server cannot be restored when in use. You may need to stop Exchange to prevent users and processes from connecting and using the database during restore. In detail, the steps for Recovery Scenario 2 are:

1. Stop production Exchange services (POP3, IMAP4, etc.), and Information Store
2. Flush file system buffers on production system and remote system
3. Perform FlashCopy restore from copy (B) to copy (A)
4. Resync production filesystem (i.e. make production system aware of new disks)
5. (optional) Check backup database integrity, consistency, and signatures at production server
6. Restart production Exchange services (POP3, IMAP4, etc.), and Information Store

When performing Step 3, the FlashCopy restore, be sure to restore all of the relevant volumes, especially the EDB and STM data files, all of the backed up transaction logs, and the checkpoint file.

Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

7C Recovery Scenario 3 - FlashCopy Roll-Forward Recovery

While Recovery Scenario 2's ability to recover to the time of the last backup is useful, many customers require the ability to recover to the moment of failure. Recovery Scenario 3 provides this capability, while maintaining the fast restores. Logically, the FlashCopy is from copy (B) to copy (A), shown in Figure 7 above. The individual steps are as follows:

1. Stop production Exchange services (POP3, IMAP4, etc.), and Information Store
2. Copy the transaction logs to a safe location
3. Flush file system buffers on production system and remote system
4. Perform FlashCopy restore data files only from copy (B) to copy (A)
5. Resync production filesystem (i.e. make production system aware of new disks)
6. (optional) Check backup database integrity, consistency, and signatures at production server
7. Ensure Exx.log file exists - if not, rename highest numbered log to Exx.log
8. Delete checkpoint file (.chk)
9. Restart production Exchange services (POP3, IMAP4, etc.), and Information Store
10. Exchange Server will automatically roll forward

At the completion of these steps, Exchange will have recovered up to the point of failure. As long as no transaction logs are damaged, no data is lost.

8 Process Automation and Solution Integration

In the lab, we developed various applications and utilities to fully integrate Exchange Server and ESS features to create a complete backup and recovery solution. We also created scripts to fully automate the snapshot backup and recovery scenarios described in previous sections. The scripts execute a customized set of snapshot task routines set up for a particular customer environment. The customer can use these scripts, applications and utilities as a one-stop backup and restore solution for even the most complex storage configurations and customer requirements. The solutions developed are focused on providing a lightning quick, seamless backup with little or no impact on the production environment, while also protecting against disasters using the snapshot technology of the ESS.

These solutions are designed to work in many different environments, and is compatible with various customer-preferred enterprise solutions management consoles like TME, OpenView or BMC. In establishing this snapshot backup and recovery solution with Microsoft Exchange Server and ESS, IBM has created end-to-end procedures that will enable this solution to seamlessly integrate into customer environments.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

Appendix A: Hardware and Software Configuration

Production Server:

IBM Netfinity 8500R 5RY
8-way 550 MHz Pentium III Xeon processors w/ 1MB L2 cache each
2 GB memory
2 9.1 GB 7200 RPM internal SCSI disk drive
2 Emulex 8000 Fibre Channel adapter (1 Gbit)
1 IBM Netfinity 10/100 Ethernet adapter
Windows 2000 Advanced Server, Service Pack 2
Exchange Server 2000 Enterprise Edition, Service Pack 1

Secondary Server:

IBM Netfinity 7000 M10
4-way 500 MHz Pentium III processors
1 GB memory
2 9.1 GB 7200 RPM internal SCSI disk drive
2 Emulex 8000 Fibre Channel adapter (1 Gbit)
1 IBM Netfinity 10/100 Ethernet adapter
Windows 2000 Advanced Server, Service Pack 2
Exchange Server 2000 Enterprise Edition, Service Pack 1

Production Storage Server:

IBM Enterprise Storage Server F20
16 GB cache
64 18.2 GB 10000rpm HDDs
12 Fibre Channel adapters
4 dual port ESCON adapters

Backup Storage Server:

IBM Enterprise Storage Server F20
16 GB cache
64 18.2 GB 10000rpm HDDs
12 Fibre Channel adapters
4 dual port ESCON adapters

Fibre Channel Switch:

McData ED-5000 Fibre Channel Director
32 ports



Appendix B: Operating System Considerations

Snapshot disk-copying technologies have only recently gained popularity in the marketplace. Most operating systems predate these technologies, and as such, these operating systems were not designed to accommodate them. Therefore, until these operating systems fully integrate these advanced technologies, there are some aspects to consider when using FlashCopy and PPRC to ensure a smooth backup and restore. These aspects are related to volume management in Windows only, and not to Exchange Server. For more information on using FlashCopy and PPRC in a Windows environment, please consult the IBM Redbook, *Implementing ESS Copy Services on UNIX and Windows NT/2000*, available at <http://www.redbooks.ibm.com>. A few of the operational considerations affecting Windows are briefly outlined below:

Windows NT considerations

Both PPRC and FlashCopy are supported when using simple disks and fault-tolerant disks (such as volume sets). Consider these four tips when planning for Copy Services on Windows NT:

- For fault-tolerant disks, essential configuration information is stored in the Windows Registry (not on the actual disk). Therefore, when initially defining volume sets, (since FlashCopy and PPRC does not copy the data stored in the Windows Registry) a utility from IBM called FCVolSet automatically performs this registry copy. Alternatively, a special procedure can be used, as outlined below:
 1. Define and format on primary system
 2. Document the Drive Signature, Partition Number, and Order Selected
 3. Perform FlashCopy
 4. Run FTEdit (available on the NT resource kit) on secondary and define volume set
 5. Repeat whenever volume set is changed
- After Service Pack 6, it is possible to have the FlashCopy source and target volumes accessible by the same server. In this case, use Disk Administrator to write a different disk signature on the target volume and assign a drive letter. Prior to Service Pack 6, the FlashCopy source and target volumes must be attached to different servers.
- To avoid rebooting after a FlashCopy, define an identical set of disks on the target machine. Then, use Disk Administrator to unassign the drive letter, perform the FlashCopy and reassign the drive letter. After performing these steps, the drive is immediately available.
- To flush file system buffers in Windows NT and 2000, use Disk Administrator to unmount (unassign) the drive letter. To fully automate the entire solution, you can develop a command-line utility to perform the file system flush without unassigning the drive letter, as we did in the lab.

Windows 2000 considerations

Windows 2000 supports two types of disks: basic disks and dynamic disks. Basic disks are the same as Windows NT disks with the same restrictions. For dynamic disks, Windows 2000 incorporates a volume manager called the Logical Disk Manager (LDM). The LDM can create five types of dynamic volumes: simple, spanned, mirrored, striped, and RAID-5.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

On Windows NT the information relating to the disks is stored in the Windows NT registry. With Windows 2000, this information is stored on the disk drive itself in a partition called the LDM database, which is kept on the last few tracks of the disk. Each volume has its own 128 Bit Globally Unique Identifier (GUID). As the LDM database is stored on the physical drive itself, with Windows 2000 it is possible to move disk drives between different computers.

Having the drive information stored on the disk itself imposes some limitations when using Copy Services functionality on a Windows 2000 system with dynamic disks:

- The source and target volumes must be of the same physical size. Normally the target volume can be bigger than the source volume. With Windows 2000 this is not the case, for two reasons:
 1. The LDM database holds information relating to the size of the volume. As this is copied from the source to the target, if the target volume is a different size from the source, then the database information will be incorrect, and the host system will return an exception.
 2. The LDM database is stored at the end of the volume. The copy process is a track-by-track copy, and unless the target is an identical size to the source, the LDM database will not be at the end of the target volume.
- It is not possible to have the source and target FlashCopy volumes on the same Windows 2000 system. Each dynamic volume has its own 128 Bit Globally Unique Identifier (GUID). As its name implies, the GUID is unique to one volume on one system. When performing a FlashCopy, the GUID is copied as well, so this means that if you try to mount the source and target volume on the same host system, you would have two volumes with exactly the same GUID. This is not allowed, and you will not be able to mount the target volume.
- Each disk contains information about every other dynamic disk on the system. Therefore, after a FlashCopy or PPRC, the information on the other disks may be rendered inaccurate. Windows only checks the information in the LDM database on bootup. While the disks can be used without rebooting, as in Windows NT, the drives will continue to work as expected only until the first reboot. The **first** reboot after a FlashCopy will require the use of Disk Management to “import” the volume. Therefore, we recommend rebooting the system first and performing an import before using the disk.



Storage Solutions for Microsoft Exchange Server 2000: Snapshot Backup and Recovery with the IBM TotalStorage™ Enterprise Storage Server

References

1. "Microsoft Exchange Server 2000 Server Database Recovery" -- White Paper
<http://support.microsoft.com/support/exch2000/whitepapers/e2kwps.asp>
2. "Storage Management for SAP and DB2 UDB: Split Mirror Backup / Recovery with IBM's Enterprise Storage Server (ESS)" -- White Paper
<http://www.storage.ibm.com/hardsoft/diskdrls/technology.htm>
3. "Planning Your Implementation of Microsoft Exchange Server" -- White Paper
<http://www.microsoft.com/exchange/techinfo/planning/2000/Planning.asp>
4. "SAP R/3 Storage Management Split Mirror Backup & Recovery on IBM's Enterprise Storage Server on DB2 OS390" -- White Paper <http://www.storage.ibm.com/hardsoft/diskdrls/technology.htm>
5. *Implementing ESS Copy Services on UNIX and Windows NT/2000* -- IBM Redbook, Document Number SG24-5757-00, March 2001
6. *IBM Enterprise Storage Server* -- IBM Redbook Document Number SG24-5465-00, July 1999
7. "Microsoft Exchange Server 2000 Internals: Quick Tuning Guide" -- White Paper
<http://www.microsoft.com/Exchange/techinfo/administration/2000/E2KTuning.asp>
8. *IBM Enterprise Storage Server Performance Monitoring and Tuning Guide* -- IBM Redbook, Document Number SG24-5656-00, March 2000
9. *Using Tivoli Data Protection for Microsoft Exchange Server* -- IBM Redbook, Document Number SG24-6147-00, May 2001, <http://www.redbooks.ibm.com>

