

June 2002  
16QT-0502A-WWEN

Prepared by Network Storage  
Solutions, Industry Standard Server  
Solutions and Strategies, and HP  
Services

Compaq Computer Corporation

## Contents

<b>1.0</b>	<b>Understanding Hot Split Replicated Volumes</b> .....	<b>3</b>
<b>2.0</b>	<b>Online and Offline Techniques</b> .....	<b>3</b>
2.1	Evaluating Options .....	3
2.2	Benefits of Using Online Replicated Volumes.....	4
<b>3.0</b>	<b>What You Need to Know</b> .....	<b>4</b>
3.1	Process Overview.....	4
3.2	Managing Database and Transaction Log Files .....	5
3.3	Recovering Exchange 2000 Information Stores .....	6
<b>4.0</b>	<b>Summary of Best Practices</b> .....	<b>6</b>
<b>5.0</b>	<b>Risks of Using Online Replicated Volumes</b> .....	<b>8</b>
<b>6.0</b>	<b>Working with HP Services</b> .....	<b>8</b>
<b>7.0</b>	<b>Related Information</b> .....	<b>9</b>
7.1	Related Documents .....	9
7.2	Related Websites .....	10
<b>8.0</b>	<b>HSG80 Test Case</b> .....	<b>11</b>
8.1	HSG80 Test Environment... ..	11
8.2	Configuration Settings .....	11
8.3	Testing Hot Split Replication in a Non-clustered Environment .....	12
8.4	Testing Hot Split Replication in a Clustered Environment....	22
8.5	Sample Output from ESEFILE and ESEUTIL.....	27
8.6	Sample DumpCfg Output ...	30

# Hot Split Replicated Volumes for Microsoft Exchange 2000 Backup and Recovery

## *Abstract*

This paper describes the viability of managing Microsoft Exchange 2000 Information Stores with hot split replicated volumes. Replicated volumes are copies of the database, also known as clones or snapshots. Hot split refers to the process of creating volumes without disrupting service to end users.

For businesses that require 24 x 7 access to email, hot split replicated volumes can successfully be used to recover from disasters and to meet Service Level Agreements. Hot split replicated volumes accelerate disaster recovery while limiting the downtime typically associated with the traditional offline methods.

The intended audience is made up of experienced Exchange administrators, who understand how to work with clones and snapshots using array controllers in clustered and non-clustered environments.

Using this information as a guide, customers can work with HP Services to prepare and test their own recovery methods using hot split replicated volumes. By engaging the expertise of HP Services from design through support, customers can be assured of a successful end-to-end implementation.

©2002 Compaq Information Technologies Group, L.P.

Compaq, the Compaq logo, StorageWorks, and SANworks are trademarks of Compaq Information Technologies Group, L.P., in the U.S. and/or other countries.

Microsoft, MS-DOS, Windows, and Windows NT are trademarks of Microsoft Corporation in the U.S. and/or other countries.

All other product names mentioned herein may be trademarks in their respective countries.

Hot Split Replicated Volumes for Microsoft Exchange 2000 Backup and Recovery  
Technical Paper prepared by Network Storage Solutions, Industry Standard Server Solutions and  
Strategies, and HP Services.

Second Edition June 2002  
Part Number 16QT-0502A-WWEN

## 1.0 Understanding Hot Split Replicated Volumes

Popular backup solutions include using clones, which are physical copies of data, to perform a rapid restoration of application environments. Clones are also known as Business Continuance Volumes (BCVs). Other solutions involve using snapshot technology to recover database environments in less time than with conventional tape restore methods.

Now you can also use online replicated volumes for Microsoft Exchange 2000 backup and recovery in conjunction with traditional backups to tape or disk media. The process of breaking off a clone to create an online replicated volume is called “hot split” replication. The term “hot split” means that the Exchange Information Stores are online and mounted when the mirror set is broken. When you create an online replicated volume, you are creating a copy of the database while it is running.

With the Exchange 2000 Information Stores online, you break off a physical copy of a volume, without disrupting the application. By manipulating log files, you ensure point-in-time data integrity. After the clone is created, you have replicated volumes of all the databases and can swap these volumes with production volumes in case of failure. Exchange mounts the replicated volume as if it were the original disk. Log files are replayed to ensure that the database is brought to a current state and that no data is lost.

Exchange administrators greatly accelerate recovery from database problems when using hot split replicated volumes. It is important, though, to keep in mind that using hot split replicated volumes bypasses the Exchange backup API that Microsoft provides. Microsoft acknowledges that hot split backups are successful although these backups depend on the following:

- Specialized hardware that can create clones or replicated volumes.
- Methods that reliably capture Exchange database files in a “soft crash equivalent” state.
- Ability to validate the replicated volumes without damaging the files or file system.

It is critically important that you establish and maintain a support relationship with HP Services to ensure successful design, implementation, and ongoing supportability when deploying this methodology in your environment.

## 2.0 Online and Offline Techniques

In the past, clones and snapshots were recommended for offline storage management. Stores that were being replicated were dismounted or the Exchange Service was stopped. Both of these procedures briefly disrupted service to end users, which can be unacceptable in 24 x 7 environments. To eliminate the disruption of service, you can implement an online replicated volume methodology, which creates copies of the databases without dismounting the databases or Storage Groups.

### 2.1 Evaluating Options

For Exchange 2000 customers, HP recommends the following methods for backup and recovery:

- If the business can tolerate minutes of downtime, offline split methodology, which includes implementing snapshot and clone technology, is the most reliable.

- If 24 x 7 access is a requirement, HP has tested and validated the use of online replicated volumes to create copies and recover Exchange 2000 Information Stores.
- Before you implement the use of hot split replicated volumes, it's important that you thoroughly understand the methodology, evaluate the risks, and partner with HP Services. As Microsoft's only "Worldwide Prime Integrator for Exchange 2000," and the first company to break the 2 million seat mark for Exchange 2000 implementations, HP Services can leverage its unsurpassed experience directly into a solution tailored to your environment.

## 2.2 Benefits of Using Online Replicated Volumes

For end users in 24 x 7 environments, implementing the hot split methodology provides the most important benefit — access to email at all times.

For Exchange administrators, this implementation provides the following benefits:

- Confidence that HP has tested and validated the methodology, which provides uninterrupted access to email.
- Access to HP service professionals who design, implement, and support worldwide Exchange 2000 environments.
- The ability to create replicated volumes without dismounting Exchange 2000 Storage Groups, which does the following:
  - Eliminates potential disruption to end users.
  - Supports the requirements of 24 x 7 environments.
- No additional software or hardware needed to support your existing clone technology.
- Implementation has been tested according to Microsoft guidelines.
- Discussion of risks, including managing database and transaction log files, provides the information needed for appropriate evaluation.

## 3.0 What You Need to Know

This section covers the following topics:

- Process Overview
- Managing Database and Transaction Log Files
- Recovering Exchange 2000 Information Stores

### 3.1 Process Overview

Here is an overview of the preparation and recovery processes you can use to test the hot split solution in your Exchange environment:

- Create the replicated volumes
- Manage the database and log files

- Recover from failures using the replicated volumes
- Make informed management decisions concerning risk

The general success criteria for recovery include a mounted database without any loss of data. HP advises customers to develop an acceptance test plan with specific acceptance criteria that meets their exact requirements. The customer along with qualified HP personnel can then execute the plan.

Before you implement the hot split solution, it's imperative that you thoroughly understand the processes and become familiar with the tasks necessary to recover an Information Store. You can use hot split replicated volumes in non-clustered and clustered environments. And you can use the following arrays:

- *StorageWorks* by Compaq HSG80 array controller and *SANworks* by Compaq Enterprise Volume Manager (EVM). Although EVM is the preferred method, you can also use *SANworks* by Compaq Command Scriptor to manually split the clone.
- *SANworks* by Compaq EVA/HSV110 array controller and the HSV Element Manager. Use the Storage System Scripting Utility V2.0 (SSSU) for Enterprise Virtual Array to write scripts that will split the clone.
- HP Surestore Virtual Array 7400, HP CommandView SDM, and HP Surestore Business Copy VA. The VA 7400 array supports more than 1000 snapshots in both a clustered and non-clustered environment.
- HP Surestore Disk Array xp512/xp48, HP Surestore Raid Manager XP and HP Surestore Business Copy XP. The xp512/xp48 array support a maximum of 9 mirror copies in both a clustered or non-clustered environment.

### 3.2 Managing Database and Transaction Log Files

When you implement hot split technology, it's important to keep in mind that cloning replaces all the databases on the same volume at the same time. Therefore you need to carefully manage the log files to ensure data integrity. It's critical that you familiarize yourself with the guidelines described in Microsoft's Knowledge Base Articles:

“[XADM: Understanding Offline and Snapshot Backups \(Q237767\)](http://support.microsoft.com/default.aspx?scid=kb;en-us;Q237767)” is located at <http://support.microsoft.com/default.aspx?scid=kb;en-us;Q237767>. It features backup solutions for Exchange, including emergency offline backups of Exchange data.

“[XADM: Hot Split Snapshot Backups of Exchange \(Q311898\)](http://support.microsoft.com/default.aspx?scid=kb;en-us;Q311898)” is located at <http://support.microsoft.com/default.aspx?scid=kb;en-us;Q311898>. It describes hot split backup solutions and snapshot solutions for Exchange data.

Here is a review of important points to keep in mind.

- When you use a hot split volume as the basis for recovery, the checkpoint file is removed from the volume, which forces all log files to replay. This procedure needs to be repeated for each Storage Group that is being recovered. The number of log files replayed determines the duration of the log replay. Each time you successfully perform a full online backup of the databases, log files are purged from disk.

Typically the administrator performs online nightly backups using tools that call the Exchange backup API. Once the backup is completed, the clone can be split off. If

any of the logs are missing, the clone cannot be used for recovery. You can determine the required log files by running `eseutil /ml`, which displays the starting log file.

- Avoid circular logging with disk cloning since the consequences are not suitable for a reliable production environment. When circular logging is enabled for an Exchange Storage Group, the Store marks a log file for reuse after its transactions are committed to the database. With circular logging disabled, log files represent the complete set of transactions that occurred since the last full backup.
- Recover the database files (EDB and STM pair) as a paired set and keep them on the same logical volume.
- Since managing the underlying clone disk will be done at the controller level, all files on the logical volumes will be affected. For ease of use and safety, it is advised that only one logical volume be created per disk array to be presented to the host. See Figure 1 for an example of logical volume I, using all of Disk 5.
- Replay all transaction log files to bring the Stores to consistency. Since disk cloning is performed on a volume level, all databases or Stores on that disk are recovered at the same time. Even if a single database of a potential five per Storage Group experiences a problem, the disk clone replaces all Stores as long as they are on the same LUN. Keep this point in mind as you design your disk layout.

### 3.3 Recovering Exchange 2000 Information Stores

During normal operations, an Exchange 2000 Information Store is considered “inconsistent” because transactions are in the process of being committed to the databases. When the Exchange Server is shut down gracefully, the database is brought to a consistent state as the services are stopped. However, when an Exchange Server experiences an interruption to service, such as loss of power, the Information Store is left in an inconsistent state.

When you start up the Information Store service, a “soft recovery” is performed as incomplete transactions are undone or redone, and new transactions are committed to bring the database to a consistent state. The server then resumes servicing messaging operations.

Soft recovery is the process of bringing the database to a consistent state after a restore procedure, using the backup, log files, and patch files.<sup>1</sup> The term “soft recovery” is used in contrast with “hard recovery,” which means the operating system or hardware suddenly becomes unavailable or fails.

Additional information about Exchange 2000 disaster recovery can be found in the paper “[Exchange 2000 Disaster Recovery for Service Providers](http://activeanswers.compaq.com/aa_downloads/6/100/225/1/42317.pdf),” available at [http://activeanswers.compaq.com/aa\\_downloads/6/100/225/1/42317.pdf](http://activeanswers.compaq.com/aa_downloads/6/100/225/1/42317.pdf).

## 4.0 Summary of Best Practices

For Exchange administrators, these best practices have been identified during the course of testing the online replication of volumes for managing 24 x 7 Exchange 2000 environments:

- Complement the traditional online backup and recovery procedures with online replication methods, such as clones and snapshots.

---

<sup>1</sup> According to Microsoft Knowledge Base article Q304125, soft recovery will succeed with little risk to the Information Store databases if there are no more than three or four outstanding log files at any point in time. However, recovering to a hot split clone involves removing the checkpoint file so that all log files on disk are replayed.

- When deploying hot split replication techniques, keep the outstanding log files to about three or four. Soft recovery succeeds with little risk to the Exchange Information Store databases when there are about three or four outstanding log files at any point in time. Microsoft defines a “soft crash” as one that occurs when a process stops unexpectedly, leaving the operating system and hardware to continue functioning normally. Microsoft estimates that there is less than a 1 in 1,000 chance that an Exchange 2000 database will not start after a “soft crash.”
- When implementing hot split replication methods, follow the guidelines documented in the Microsoft Knowledge Base articles:
  - “[XADM: Understanding Offline and Snapshot Backups \(Q237767\)](http://support.microsoft.com/default.aspx?scid=kb;en-us:Q237767)” is located at <http://support.microsoft.com/default.aspx?scid=kb;en-us:Q237767>. It features backup solutions for Exchange, including emergency offline backups of Exchange data.
  - “[XADM: Hot Split Snapshot Backups of Exchange \(Q311898\)](http://support.microsoft.com/default.aspx?scid=kb;en-us:Q311898)” is located at <http://support.microsoft.com/default.aspx?scid=kb;en-us:Q311898>. It describes hot split backup solutions and snapshot solutions for Exchange data.
- Avoid using “circular logging.” When circular logging is enabled for an Exchange Storage Group, the Store marks a log file for reuse after its transactions are committed to the database. With circular logging disabled, log files represent the complete set of transactions that occurred since the last full backup.
- When designing your Store disk layout, it is important to consider a design that facilitates replaying log files to ensure the point-in-time data integrity of the Exchange Information Store. Replaying *all* transaction log files brings the Stores to consistency.
- Consider the number of users per Storage Group as the sizing determination for log file replay. If the transaction logs created during the day mean that log file replay will exceed the acceptable recovery window, then an additional Storage Group should be created and placed on a different disk volume.
- To minimize recovery time, create replicated volumes after a full backup has successfully completed, since transaction logs will be purged. This reduces the set of transaction log files to include only the incoming transactions since the full backup.
- Exchange database components such as the edb, stm, checkpoint, and log files should be the only components on replicated volumes. Other Exchange components such as the Message Transfer Agent, SMTP drop directories, and Message Tracking Logs should be placed on other volumes to simplify recovery.
- When using hot split replication in a clustered environment, it is advisable to make note of the original Disk Signatures for the Cluster Disk Resource recorded in the Registry. This ensures that DumpCfg is run against the correct disk.
- By design, Exchange 2000 Server is self-tuning and requires little or no administrative intervention to optimize performance. Let Exchange 2000 manage performance values to reduce the risk of data loss.

## 5.0 Risks of Using Online Replicated Volumes

This section describes the risks you need to keep in mind as you evaluate the hot split technique for your business.

- Hot split replicated volumes bypass the Microsoft online backup API to back up and restore Exchange data. Microsoft does not test, certify, or debug scripts and procedures for hot split backups of Exchange data files. When using hot split methods, Microsoft Product Support Services (PSS) can help customers to analyze transaction log files sets, although the backup solution vendor or Exchange Administrator is responsible for ensuring the integrity of all the files at each stage of the process.
- Hot split replicated volumes generate an inconsistent database by definition. Replaying log files is required to roll forward any outstanding transactions. Database integrity checking with utilities such as ESEUTIL and ESEFILE, will not guarantee the integrity of the database.
- Under stress testing, hot split replicated volumes resulted in database corruption when the Exchange Server was under heavy load and, concurrently, the Store process was actively allocating memory. Therefore, it is inadvisable to perform hot split replications under these simultaneous conditions.

Microsoft acknowledges that the success of a hot split backup solution depends on carefully designed hardware, methods that reliably capture Exchange database files in a “soft crash equivalent” state, and validation of the Exchange database capture without damage to the files or file system. **Customers can benefit by working with HP Services to ensure the design and supportability of their solution, and ultimately leverage the safeguards necessary for successful implementation.**

## 6.0 Working with HP Services

You can work with HP Services to prepare and test your own recovery methods using hot split replicated volumes. In addition to helping you design a recovery strategy, HP has a full spectrum of services available for you to think about. Some of these include evaluating your environment, preparing pilot programs for review, and designing and implementing customized Exchange environments. Here is a list of services for you to review:

- As you prepare your existing infrastructure for a Microsoft Exchange 2000 implementation, “[ExpertCheck for Exchange 2000](http://www.compaq.com/services/messaging/mg_expertcheck.html)” provides an opportunity to evaluate your environment, identifying ways to achieve a more economical and secure messaging environment. See [http://www.compaq.com/services/messaging/mg\\_expertcheck.html](http://www.compaq.com/services/messaging/mg_expertcheck.html) for more information.
- With “[Preparation Assessment for Exchange 2000](http://www.compaq.com/services/messaging/mg_expertcheck.html),” our service professionals analyze and review your current messaging strategy, business requirements, network and operating environment, and email usage. These professionals will recommend options and estimate the effort to migrate to Exchange 2000 and maximize your return on investment. For more information, see [http://www.compaq.com/services/messaging/mg\\_expertcheck.html](http://www.compaq.com/services/messaging/mg_expertcheck.html).
- As part of the complete portfolio of proven lifecycle services, “[Evaluation Pilot for Exchange 2000](#)” can be used to develop and customize applications, test new versions



and back up production environments. For more details, see [http://www.compaq.com/services/messaging/mg\\_pilot.html](http://www.compaq.com/services/messaging/mg_pilot.html).

- “[Planning & Design for Exchange 2000](http://www.compaq.com/services/messaging/mg_design.html)” provides a well designed plan for your new messaging environment, mapping Exchange 2000 capabilities to your objectives. See [http://www.compaq.com/services/messaging/mg\\_design.html](http://www.compaq.com/services/messaging/mg_design.html) for more information.
- If you’re looking to simplify the complexities of managing your messaging environment, “[Customized Management](http://www.compaq.com/services/messaging/mg_manage.html)” is tailor-made for you. See [http://www.compaq.com/services/messaging/mg\\_manage.html](http://www.compaq.com/services/messaging/mg_manage.html) for more information.
- “[Help Desk Support](http://www.compaq.com/services/messaging/mg_helpdesk.html)” provides you with expert help desk support resources for your Microsoft Exchange environment, as described at [http://www.compaq.com/services/messaging/mg\\_helpdesk.html](http://www.compaq.com/services/messaging/mg_helpdesk.html).
- If you’re just getting started with Windows 2000, see “[Deployment Services for Windows 2000](http://www.compaq.com/services/messaging/mg_helpdesk.html)” at [http://www.compaq.com/services/messaging/mg\\_helpdesk.html](http://www.compaq.com/services/messaging/mg_helpdesk.html).
- For installation and startup for the storage in your environment, see “[CarePaq Services for Storage](http://www.compaq.com/services/carepaq/us/install/cp_storage.html)” at [http://www.compaq.com/services/carepaq/us/install/cp\\_storage.html](http://www.compaq.com/services/carepaq/us/install/cp_storage.html).
- HP offers standard, high-availability, and business critical levels of support for Windows 2000 environments. Support is provided for more than 150 Windows NT, Windows 2000, and NetWare products, including SQL Server and Exchange. For more information see [http://www.compaq.com/services/windows2000/w2\\_hardware.html](http://www.compaq.com/services/windows2000/w2_hardware.html)
- HP helps you optimize availability and reduce downtime risks across six critical IT domains: management, network, application, system software, system hardware, and physical environment. Services are specifically designed for Storage Area Networks, including *Availability Assessment Service for SANs*, *Availability Monitoring Service for SANs* and *Availability Management Service for SANs*. For more information see <http://www.compaq.com/services/available/index.html>.

## 7.0 Related Information

Here is an additional list of documents and websites that you may find helpful.

### 7.1 Related Documents

- [Ultimate Windows 2000 and Exchange 2000 Services White Paper](ftp://ftp.compaq.com/pub/services/feature/w2k_e2k.pdf)  
[ftp://ftp.compaq.com/pub/services/feature/w2k\\_e2k.pdf](ftp://ftp.compaq.com/pub/services/feature/w2k_e2k.pdf).
- [Exchange 2000 Disaster Recovery for Service Providers](http://activeanswers.compaq.com/aa_downloads/6/100/225/1/42317.pdf)  
[http://activeanswers.compaq.com/aa\\_downloads/6/100/225/1/42317.pdf](http://activeanswers.compaq.com/aa_downloads/6/100/225/1/42317.pdf).
- [Microsoft Exchange 2000 Internals: Quick Tuning Guide](http://www.microsoft.com/technet/treeview/default.asp?url=/TechNet/prodtechnol/exchange/maintain/optimize/exchtune.asp)  
<http://www.microsoft.com/technet/treeview/default.asp?url=/TechNet/prodtechnol/exchange/maintain/optimize/exchtune.asp>.

## 7.2 Related Websites

- [Microsoft Services and Solutions](http://www.compaq.com/services/microsoft/index.html)  
<http://www.compaq.com/services/microsoft/index.html>.
- [Compaq VP authors technical bestseller on Exchange](http://www.compaq.com/services/feature/index_redmond.html)  
[http://www.compaq.com/services/feature/index\\_redmond.html](http://www.compaq.com/services/feature/index_redmond.html).
- [Worldwide Prime Integrator for Microsoft Exchange 2000](http://www.compaq.com/services/messaging/index.html)  
<http://www.compaq.com/services/messaging/index.html>.
- [Milestone: Compaq reaches 2 million Microsoft Exchange 2000 seats](http://www.compaq.com/newsroom/pr/2002/pr2002042302.html)  
<http://www.compaq.com/newsroom/pr/2002/pr2002042302.html>.
- [StorageWorks Enterprise Virtual Array](http://www.compaq.com/products/storageworks/enterprise/index.html)  
<http://www.compaq.com/products/storageworks/enterprise/index.html>.
- [SANworks Enterprise Volume Manager](http://www.compaq.com/products/sanworks/evm/index.html)  
<http://www.compaq.com/products/sanworks/evm/index.html>.
- [Virtualized Storage Management for Exchange 2000](http://www.compaq.com/products/storageworks/solutions/vsm-e2000/index.html)  
<http://www.compaq.com/products/storageworks/solutions/vsm-e2000/index.html>.
- [Microsoft Exchange Server Solutions](http://www.compaq.com/solutions/showroom/mses.html)  
<http://www.compaq.com/solutions/showroom/mses.html>.

## 8.0 HSG80 Test Case

This section describes a scenario for using hot split replicated volumes to backup and recover from disaster failures in a *StorageWorks* HSG80 environment. You can review the material and evaluate it as you think about using replicated volumes in your own business.

### 8.1 HSG80 Test Environment

This section details the results of validation testing performed to assess the optimum design for hot split cloning. The first environment represents a simple SAN design with a *StorageWorks* HSG80 controller servicing several Exchange servers. Both command scripts and *SANworks* EVM were used to manage the disk configuration, clone splits, and timing of commands.

The second environment is more complex, with a fully fault-tolerant design involving Microsoft Cluster Server (MSCS) and *SANworks* by Compaq Secure Path for managing multiple, fault-redundant paths to the storage. The addition of the clustered Exchange Server changes the recovery procedures, which are included.

Non-clustered and clustered environments have been tested with *StorageWorks* HSG80 and *SANworks* HSV110 array controllers. In both non-clustered and clustered environments, the processes are the same although some specific tasks are different. However, keep in mind that the test environments described in sections 8.3 and 8.4 only apply to the HSG80 array controller.

### 8.2 Configuration Settings

The following configuration parameters were tested to determine how they might impact the success of using hot split replication. While testing determined that the parameters were not absolutely necessary, these changes may reduce the risk of clone recovery failures. The configurations are categorized by operating system (Windows 2000) and Exchange Server.

#### ***Operating System (Windows 2000) Configuration***

##### **Disk Format Parameters**

Similar to the issues discussed in the above section on chunksize, the database volume was formatted with the minimum allocation unit (4K) in order to reduce the probability that a database page (4K) would be split across multiple units. It was determined that this setting was not necessary to perform successful clone splits.

##### **Reducing Outstanding I/O to Disk**

In order to reduce the amount of outstanding I/O traffic to the database files during the clone split, several methods were used to flush the NTFS file system cache. It was determined through testing that this was not necessary to perform successful clone splits.

##### ***Exchange Server Configuration***

---

**Caution:** Exchange 2000 Server is designed to be self-tuning, requiring little or no administrative intervention to optimize performance. It is recommended that Exchange 2000 manage all performance values to reduce the risk of data loss.

---

## Reducing Outstanding I/O to Disk

In order to reduce the amount of outstanding I/O traffic to the database files during the clone split, the ESE backup API was called by triggering a full backup. However it was determined through testing that this was not necessary to perform successful clone splits.

### Log Buffers

Log Buffers are the amount of memory that Exchange 2000 uses for incoming transactions before writing to transaction log files. In the “Microsoft Exchange 2000 Internals: Quick Tuning Guide,” which is located at this address:

<http://www.microsoft.com/technet/treeview/default.asp?url=/TechNet/prodtechnol/exchange/maintain/optimize/exchtune.asp>, it is mentioned that the number of default Log Buffers (84) used for Exchange 2000 Service Pack 1 and earlier is too small, and the suggested value is 500. Increasing this value improves performance, but can mean data loss if the server crashes or loses all power. It is recommended that this parameter (msExchESEParamLogBuffers attribute) be left alone for Exchange 2000 to manage, as the smaller value reduces any risk of data loss.

### Checkpoint Depth

The checkpoint depth is the number of log files whose data has not yet been committed to the database file. The current checkpoint depth is stored in the checkpoint (.CHK) file for all databases in the Storage Group (eseutil /mk E00.chk). It is possible to reduce the checkpoint depth from the default (20MB in Exchange 2000 Server Service Pack 1) for optimum cluster failover performance, as suggested in the paper “Deploying Microsoft Exchange 2000 Server Service Pack 1 Clusters.” Decreasing the checkpoint depth will increase I/O to the database volume and will have no impact on recovery to clone, as the checkpoint file will be removed and the entire transaction log set replayed.

## 8.3 Testing Hot Split Replication in a Non-clustered Environment

To validate the scripts and the ability to recover from a clone, a manual clone split with the Exchange Information Store offline was performed before the implementation of the hot split process.

With a load of 2,000 MMB2 users that was generated using LoadSim, a clone split was performed. In addition, SMTP traffic, generated using MailTest, was read via IMAP4 using Outlook Express in order to exercise the STM database files. A single Storage Group was created with two Mailbox Stores placed on the same disk. Figure 1 shows the disk layout, with Disk 5 (logical drive I:) being the database volume and Disk 6 (logical drive J:) being the log volume.

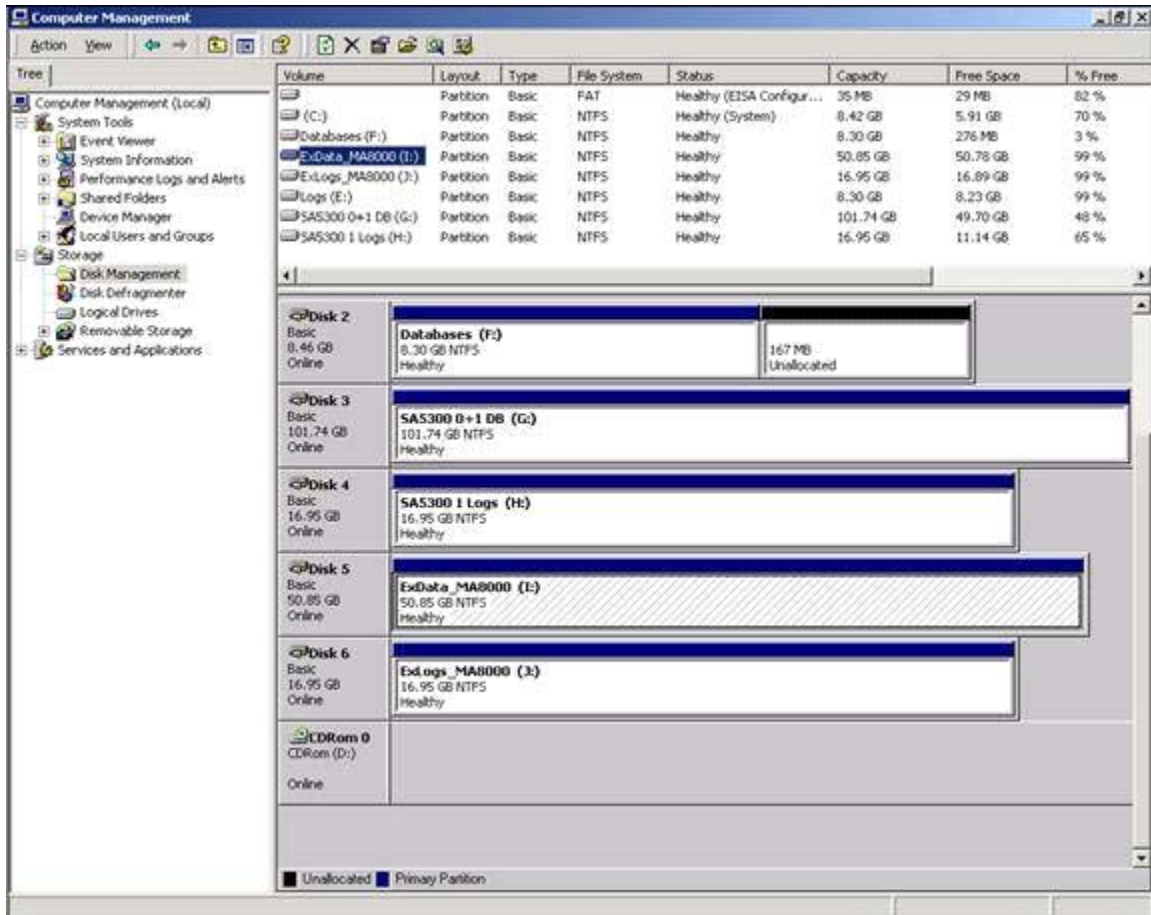


Figure 1 Disk Management view of Exchange database volume (Disk 5, Logical drive I:).

Figure 2 shows the database files after being fully populated with LoadSim email messages. Figure 3 shows the view in Exchange System Manager of the size of the mailboxes.

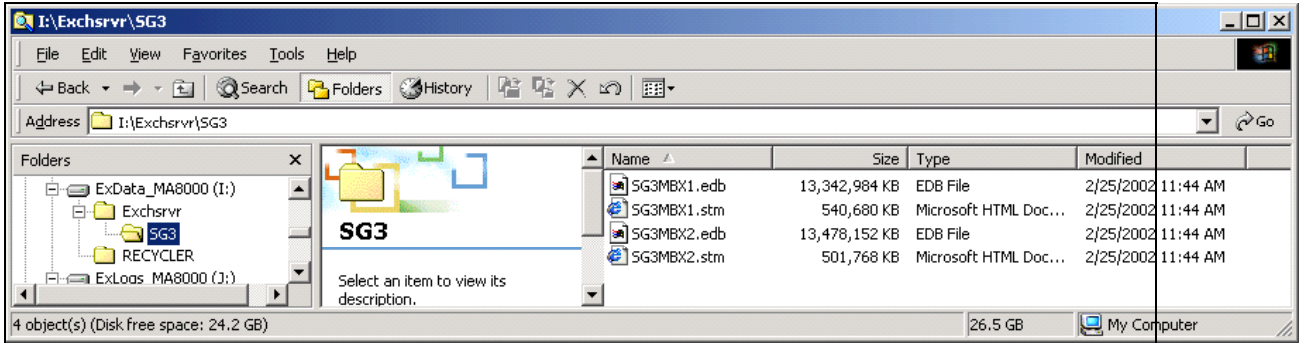


Figure 2 Storage Group 3 database files on cloned volume.

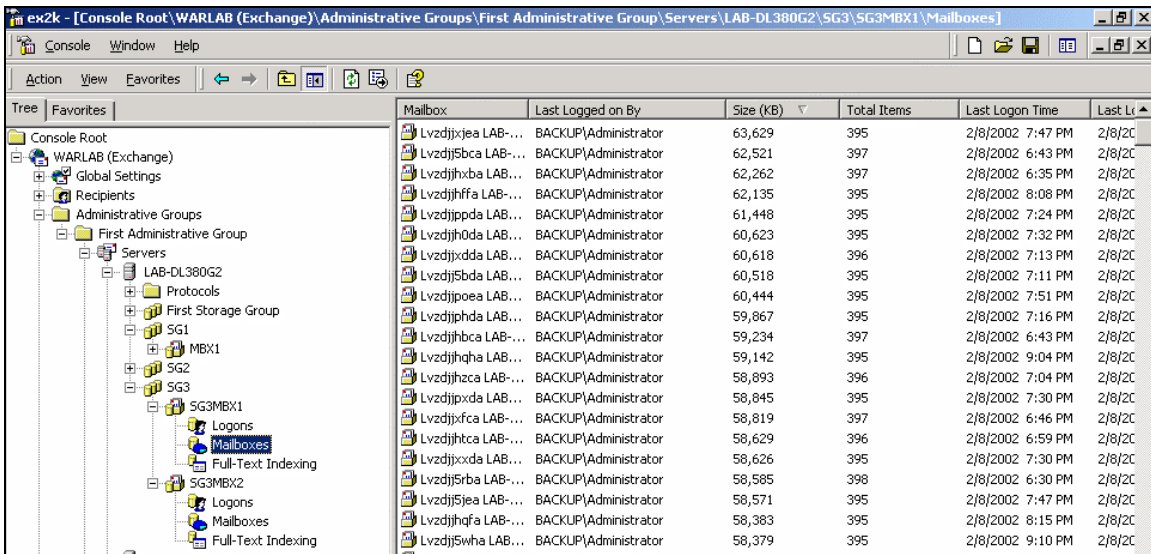


Figure 3 Size of sample mailboxes.

Figure 4 shows the database files at the time that the clone is split off from the other members of the mirror set. The files have grown from the size shown in Figure 2.

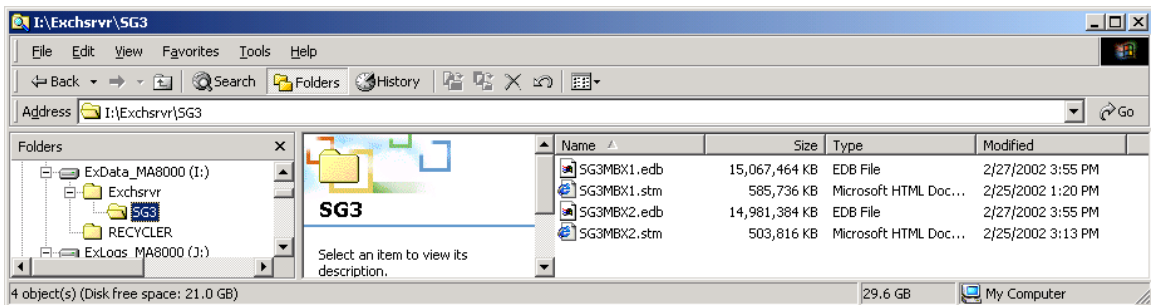


Figure 4 Database files before clone split.

### Clone Split Process

In any cloning technique, whether the Exchange Information Store is online or offline, the clone must be fully normalized before the split is performed. Table 1 shows a mirror set, M3, normalizing as the third disk is copying. You can manage this state with custom scripts or by

using *SANworks* Command Scriptor, but it is best handled by EVM as shown below. Figure 5 shows the EVM job operations, including clone creation, normalization, and split commands. In EVM, each operation or command is not called until the previous one completes successfully.

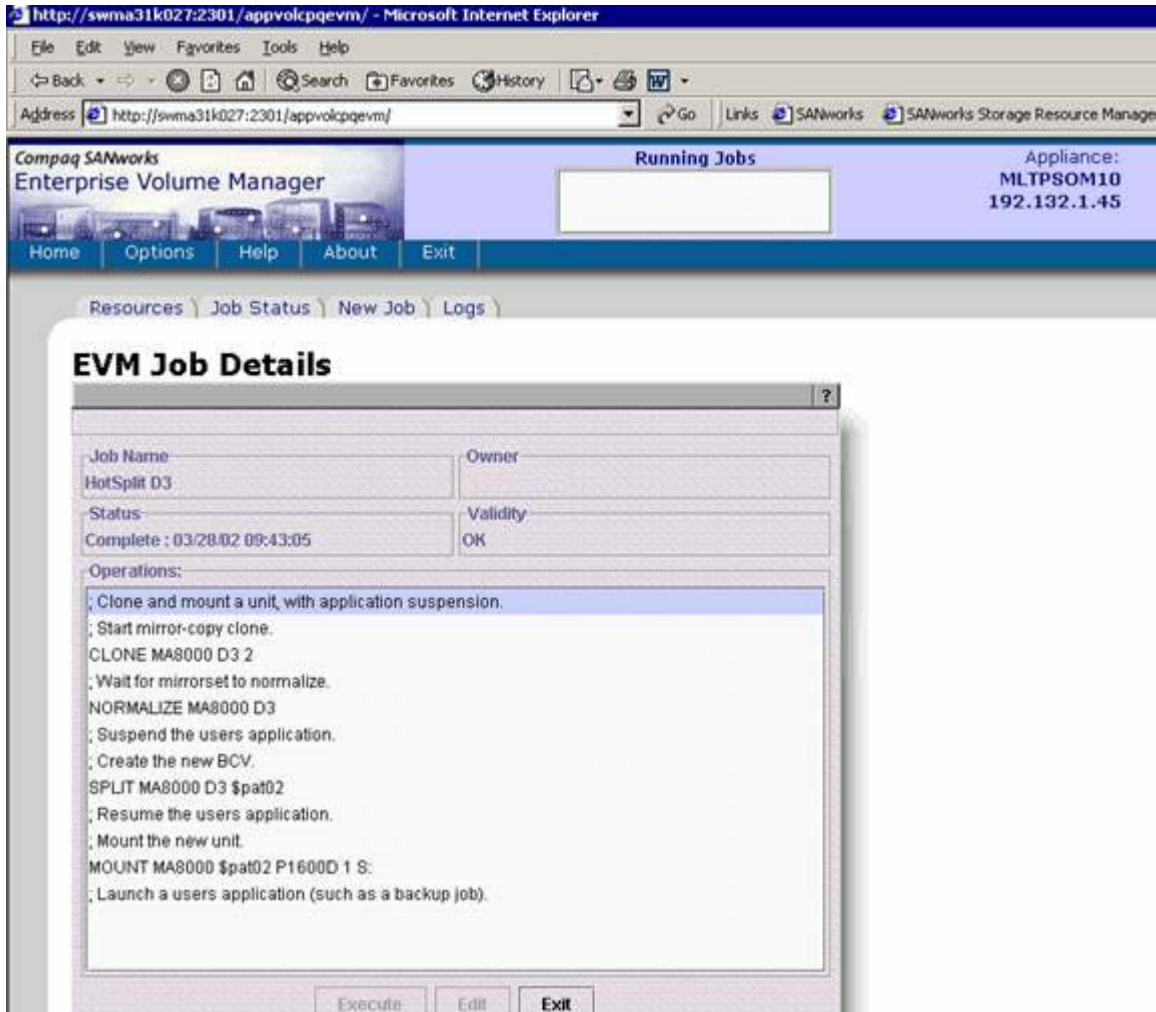
**Table 1 Output from HSG80 controller showing normalizing mirror set.**

```

HSG> sho m3
Name          Storageset          Uses          Used by
-----
-
M3            mirrorset           DISK40200     S1
                DISK50000
                DISK60300

Switches:
  POLICY (for replacement) = BEST_PERFORMANCE
  COPY (priority) = NORMAL
  READ_SOURCE = LEAST_BUSY
  MEMBERSHIP = 3, 3 members present
State:
  NORMAL
  DISK40200 (member 0) is NORMAL
  DISK60300 (member 1) is NORMAL
  DISK50000 (member 2) is COPYING      84% complete
Size:                35551389 blocks

```



**Figure 5** EVM job details showing the clone creation, normalization, and split commands.

The clone can be split manually with a direct connection, using *SANworks* Command Scripter, or with EVM. In Windows 2000, the new disk is eventually detected without a reboot. The detection process can be sped up by performing a disk rescan in Disk Manager.

### **Database Integrity Checks**

Unlike the offline backup and recovery techniques, hot split cloning generates an inconsistent database, which means that transactions are outstanding. Trying to run utilities such as ESEUTIL generates a warning stating that the database is inconsistent and that log files should be replayed before running the utility. While it is possible to run another utility, such as ESEFILE, to validate page checksums, this does not guarantee the overall integrity of the database.



### **Recovery Using a Clone**

Here is an overview of the recovery procedure to follow in a non-clustered environment, using a clone-based copy instead of a traditional online-backup set.

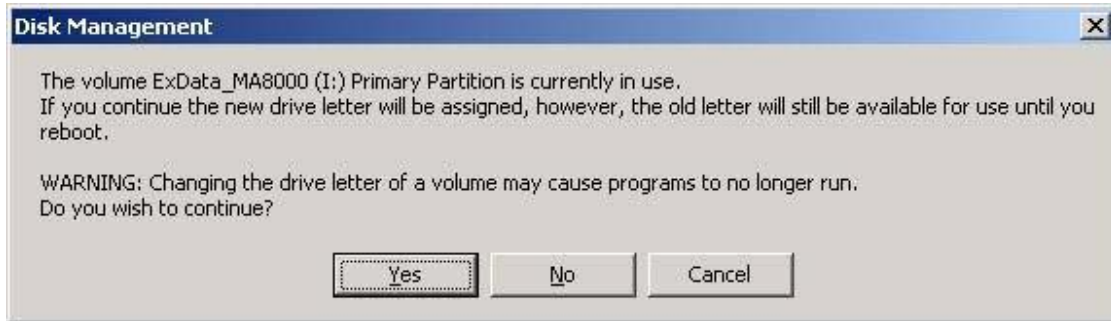
1. Stop the Microsoft Exchange System Attendant and dependent services, especially the Microsoft Exchange Information Store service.
2. In Disk Administrator, remove the drive letter of the volume where the faulty Exchange Storage Group files are located.
3. If the drive letter is locked in use as shown in Figure 6, disable the Microsoft Exchange System Attendant service and reboot the server. (Note that any services dependent on the System Attendant service will fail, such as anti-virus or backup engines.)
4. Remove the Checkpoint file for only the affected Storage Group.
5. Assign the drive letter of the Exchange Storage Group files to the clone disk.
6. Set the Stores property to allow recovery (as shown below in Figure 7).
7. Enable the Microsoft Exchange System Attendant and start all Microsoft Exchange services.
8. Watch the Application Event Log for Event 301 (Log Replay). When all the logs are recovered, verify that all Stores have mounted.
9. At this point, it's always important to back up your servers.

Here are the details involved in the process.

In Disk Manager, you need to remove the drive letter from the original volume so that it can be available for the clone. However, in order to remove the drive letter of the volume, you need to stop the Microsoft Exchange Information Store (IS) service. If there are a substantial number of outstanding transaction log files (for example, there may be other Storage Groups not involved with the clone volume about to be recovered), the process of stopping the IS service may take a few minutes.

It is *not* necessary to manually stop the Microsoft Exchange Information Store service or dismount any Store, if they will be replaced in the clone recovery process (all stores on the same disk will be replaced). This will cause an unnecessary delay as the transaction logs are committed. It may be fastest to just disable the Exchange services and reboot the server. Restarting the server shuts down the Microsoft Exchange services without advancing the checkpoint file.

Exchange expects to find the database files on the same drive letter (I: in Figure 6 below), but this drive letter may be locked in use until you restart the server. If you stop the Microsoft Exchange services and are unable to re-use the drive letter, then you need to restart the server. When the server is restarted, it is important to be able to recognize the clone and ensure that it is given the original drive letter. Make note of which disk number (in Disk Manager) the clone is before you restart.



**Figure 6** Warning when changing the drive letter for the database volume.

You must delete the checkpoint file before recovering to the clone to ensure that all log files are replayed. To prevent Exchange from starting up and attempting to mount the database, disable the Microsoft Exchange System Attendant service. Even if you do not disable the service, the Stores on the clone will not mount until you set the property for recovery as shown in Figure 7.

You may want to consider leaving the System Attendant service enabled as a valid option – it brings the server back online for all other users on the server, users who would otherwise be affected by the service interruption.

---

**Caution:** You must delete the checkpoint file before attempting to mount the Stores on the clone. If you do not delete the file and the Information Store mounts, a partial set of log files is replayed. The Stores fail to mount and will need to be restored from a backup set.

---

Figure 7 shows the Store property setting, accessed through Exchange System manager. Be sure to check this setting, which allows the database to be restored from backup. This property is stored in Active Directory so it can also be accessed using a script such as the one shown below. The `msExchPatchMDB` property needs to be set for all Stores in the affected Storage Group.

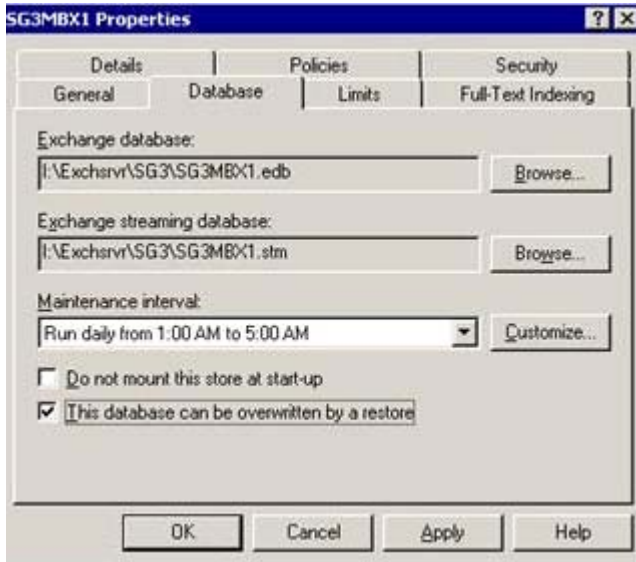


Figure 7 Store property allowing the database to be restored from backup.

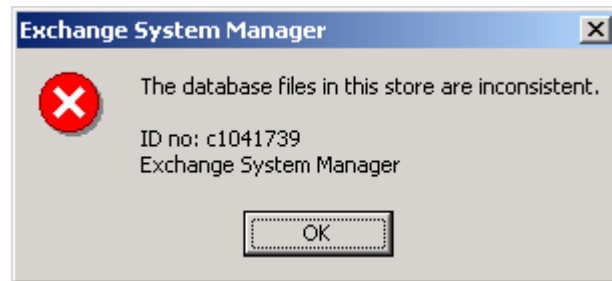
Table 2 Accessing the msExchPatchMDB property using an LDIF script.

```
C:\> ldifde -i -s domaincontroller -f patchmdbs.ldf
Contents of patchmdbs.ldf
-----
dn:
CN=SG3MBX1,CN=SG3,CN=InformationStore,CN=EX2KSERVER,CN=Servers,CN=First
Administrative Group,CN=Administrative Groups,CN=ORGNAME,CN=Microsoft
Exchange,CN=Services,CN=Configuration,DC=DOMAIN,DC=TLH
changetype: modify
replace: msExchPatchMDB
msExchPatchMDB: TRUE
-

dn:
CN=SG3MBX2,CN=SG3,CN=InformationStore,CN=EX2KSERVER,CN=Servers,CN=First
Administrative Group,CN=Administrative Groups,CN=ORGNAME,CN=Microsoft
Exchange,CN=Services,CN=Configuration,DC=DOMAIN,DC=TLH
changetype: modify
replace: msExchPatchMDB
msExchPatchMDB: TRUE
-
```

If you attempt to mount the Stores on a hot split clone volume without first setting the `msExchPatchMDB` property, you will see the message shown in Figure 8, and the Stores will fail to mount.

**Figure 8** Attempting to mount a Store without setting the `msExchPatchMDB` property.



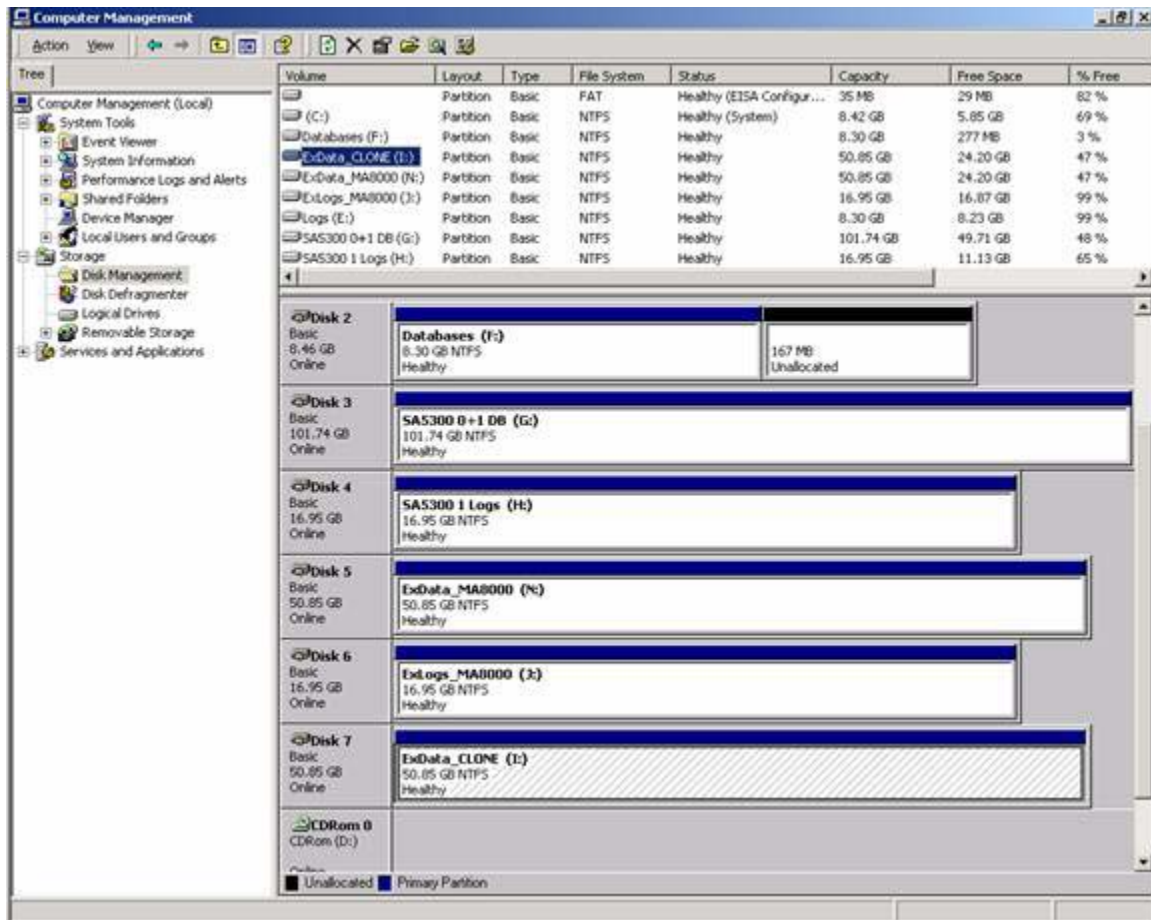


Figure 9 Disk Layout immediately after recovering to clone.

### Log File Replay

Since the transaction log checkpoint file is deleted before recovering from the clone, all transaction logs associated with the Storage Group are replayed. To minimize the recovery interval, you should create the clone only after a full backup has successfully completed and transaction log files have been purged. In this way, the set of transaction log files includes only the incoming transactions since the full backup.

Suppose the full backup completes at 4 a.m., and a hot split clone has been generated. If recovery to this clone is necessary at 4 p.m. that day, there will be 12 hours of transaction logs to replay.

The amount of transaction logs will vary per customer, depending on the email traffic. During the MMB2 LoadSim runs, the logging rate was measured to be roughly 200K per user per minute, which generated a new 5MB log file at the rate of 1.4 per minute for the 2,000 user load. This equates to 85 log files every hour, a total of 425 MB per hour.

The replay rate for these log files can vary, depending on the existing load on the Exchange Server and the RAID subsystem. In the validation environment, the log files were replayed at the rate of 16 to 20 per minute, which means that 8 hours of outstanding log files (630 to 680 log files) generated by 2,000 MMB2 users would take 32-42 minutes to replay.

At some point, if outstanding log files are allowed to accumulate, a cross-over point is reached where it is faster to restore from online backup than from the clone. This point should be calculated based on the server log replay rate and used to guide the recovery efforts.

**Note:** As a best practice, consider the number of users per Storage Group as the sizing determination for log file replay. If the transaction logs created during the day mean that log file replay will exceed the acceptable recovery window, then an additional Storage Group should be created and placed on a different disk volume.

### Hot split Clone Stress Test

Additional tests were performed in an attempt to create a clone that was not recoverable. The server was loaded with over 3,000 Users Connections using LoadSim MMB2 and also SMTP traffic at the rate of 1 message per second (to stress the .STM database).

Although it is not advisable to split a clone during a peak performance period, the testing was performed nevertheless. The tests to hot split and recover to the clone were successful even under these conditions. In fact, only one test scenario generated a clone that was not recoverable and resulted in database corruption on the clone (the original volume was still recoverable). This scenario involved hot splitting the clone while the Exchange Server was under heavy load and the Store process was still allocating memory. It is never advisable to perform hot split under these conditions in a production environment.

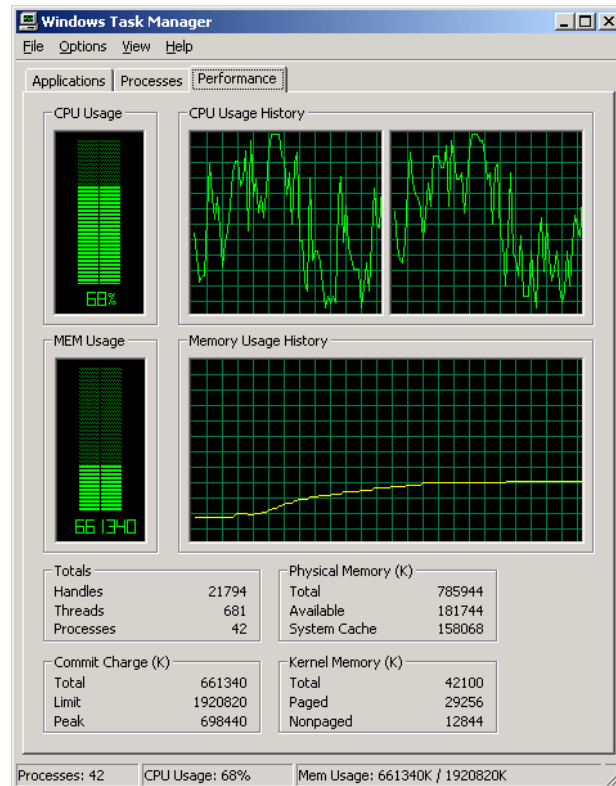


Figure 10 Performance during the stress test.

## 8.4 Testing Hot Split Replication in a Clustered Environment

The second environment was more complex, with a fully fault-tolerant design that involved Microsoft Cluster Server (MSCS) and *SANworks* Secure Path for managing multiple, fault-redundant paths to the storage.

With a load of 2,000 MMB2 users that was generated using LoadSim, a clone split was performed. A pair of HSG80 controllers serves several Exchange Servers. In addition, SMTP traffic, generated using MailTest, was read via IMAP4 using Outlook Express in order to exercise the STM database files. A Storage Group was created with two Mailbox Stores placed on the disk

to be cloned. The same procedures outlined above were followed, and all tests to create hot split replicated volumes and recover to the volume were successful.

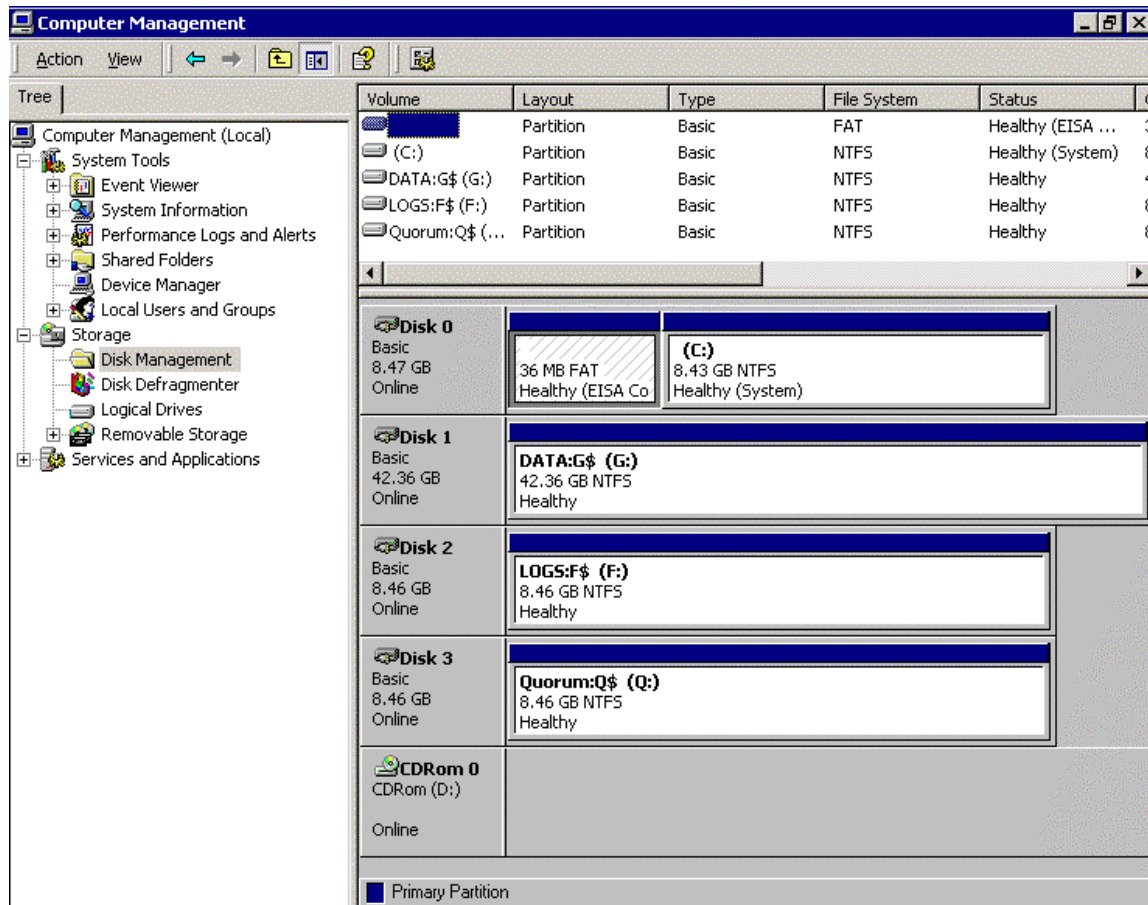


Figure 11 Disk configuration of clustered Exchange Server.

### ***Recovery Using a Clone in an Exchange Server Cluster***

Here is an overview of the recovery procedure in a clustered environment. The detailed procedures follow. The primary difference between a clustered environment and a non-clustered environment is that a cluster recognizes disk resources by signature, so the disk signatures must be modified. In addition, the second node must be considered when taking Exchange offline – both nodes need to be managed, otherwise one node will reserve access to the disk. The Cluster Administrator console is normally used to take Exchange offline, as opposed to the Services MMC.

1. Shut down the secondary node<sup>2</sup> and take the active cluster node (virtual server) offline.
2. View hidden devices in Device Manager and disable the Cluster Disk Driver (as shown in Figure 13 below). Disable the Cluster Service in the Services MMC.
3. Remove the Checkpoint file for only the affected Storage Group.
4. Reboot the server (note that any services dependent on the Cluster Service will fail).
5. Remove the drive letter of the Exchange Storage Group volume.
6. Record the disk signature to the Storage Group volume to be replaced (using DumpCfg).
7. Write a unique, temporary disk signature to the Storage Group volume to be replaced (using DumpCfg).
8. Write the disk signature noted in Step 6 to the clone volume (using DumpCfg).
9. Assign the Exchange Storage Group volume drive letter to the clone disk.
10. Set the Store(s) property to allow recovery (as detailed earlier).
11. Enable the Cluster Service in the Services MMC and the Cluster Disk Driver in Device Manager. Reboot the Server and bring the cluster node (virtual server) online.
12. Watch the Application Event Log for Event 301 (Log Replay). Verify that all Stores have mounted.

Figure 12 shows the disk configuration of the clustered Exchange Server after the clone has been split and a Disk Rescan run to detect the new disk.

---

**Caution:** You must delete the checkpoint file before attempting to mount the Stores on the clone. If you do not delete the file and the Information Store mounts, a partial set of log files is replayed. The Stores fail to mount and will need to be restored from a backup set.

---

---

**Note:** If you use Windows Terminal Server to connect to a cluster node, you may not see all of the drives in Windows Explorer. To delete the Transaction Log checkpoint file, you may need to log off and log on again or open the UNC location such as [\\server\fs](#) for the transaction log drive.

---

---

<sup>2</sup> Any failover target node in Windows 2000 Advanced Server or Datacenter Server. In Windows .NET server it may be possible to create more active nodes in a cluster. Any node that has access to the disk being recovered should be shut down to allow the swap in disks.



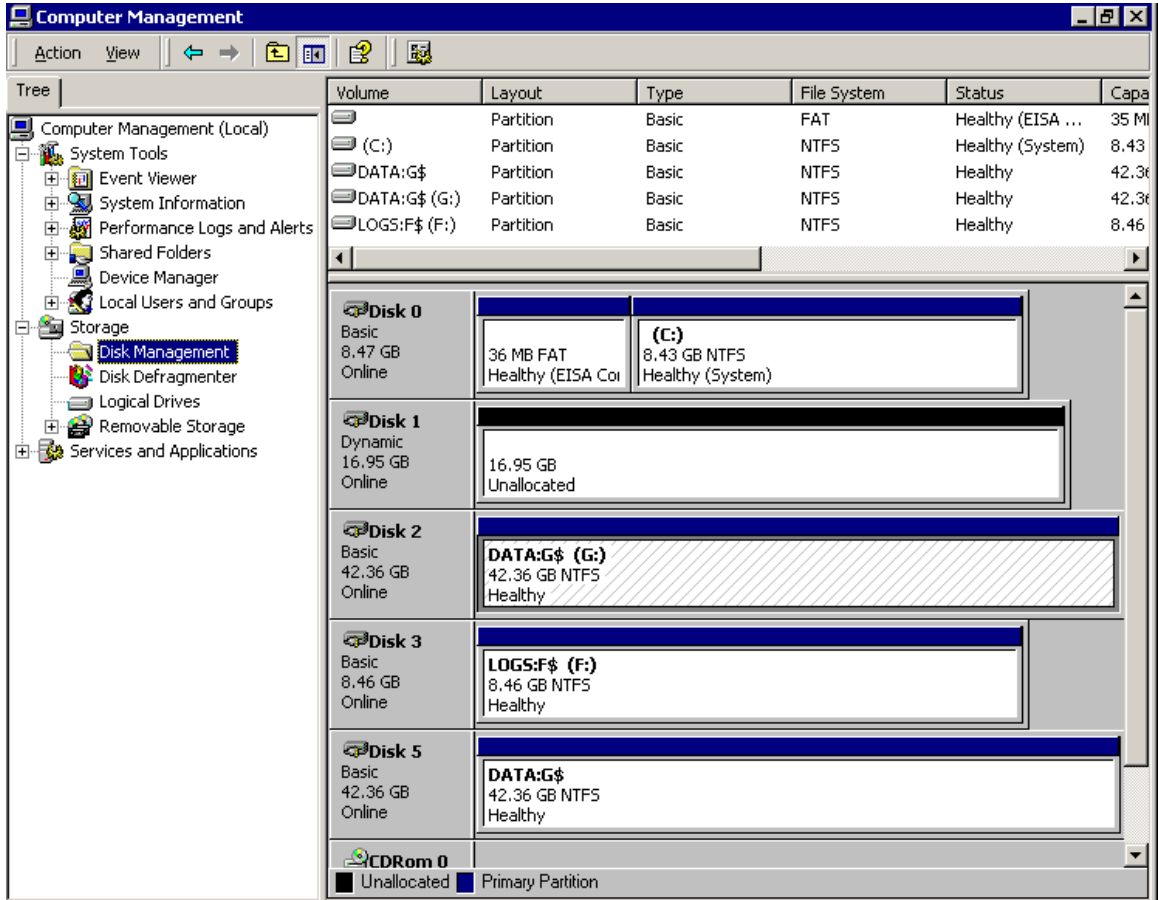


Figure 12 Disk configuration of clustered Exchange Server showing Clone (after Rescan).

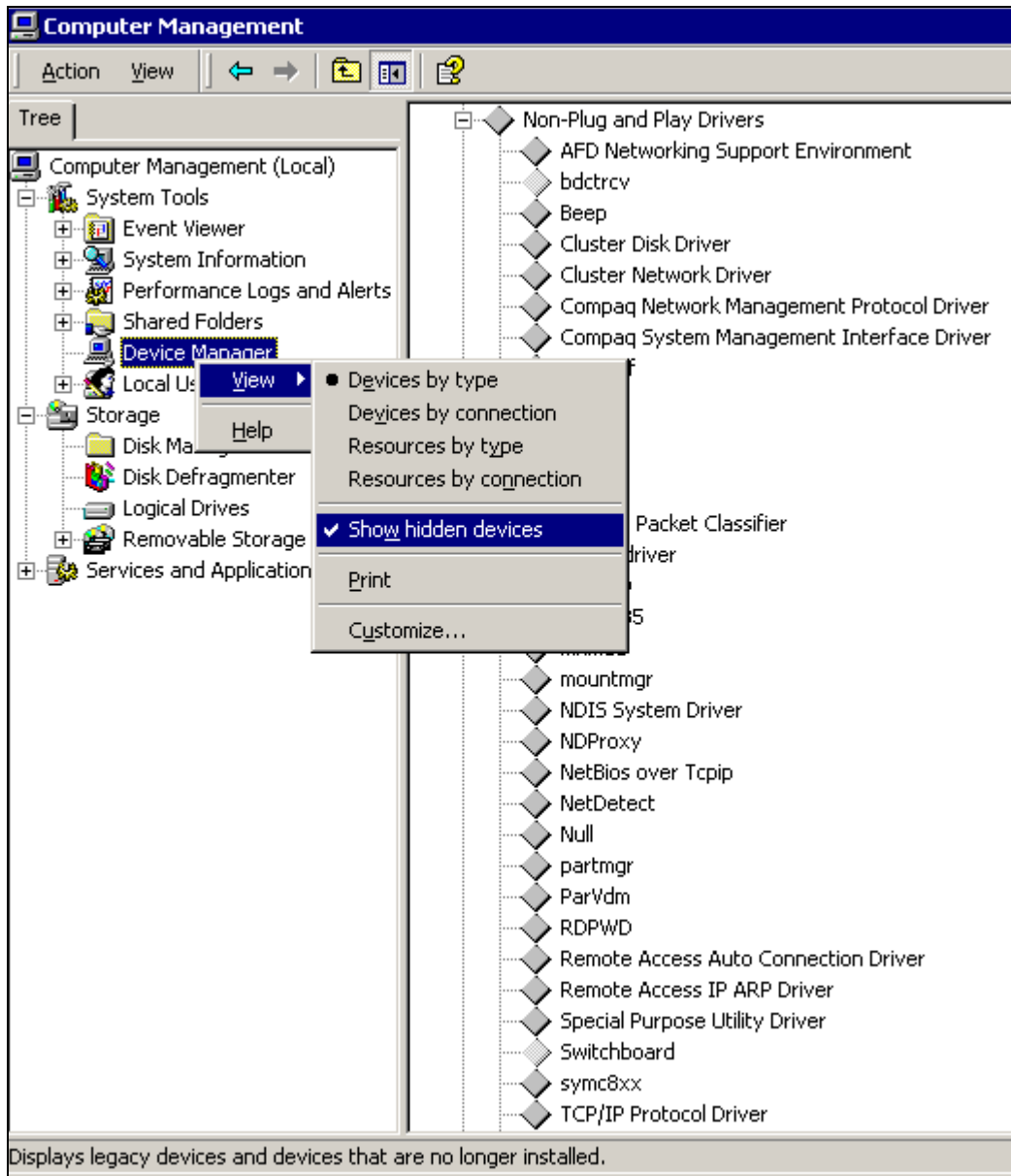


Figure 13 Disabling the Cluster Disk Driver in Device Manager.

### Cluster Disk Signatures

Cluster disk resources are recognized by disk signatures. To replace the original Exchange database volume (which is a Cluster Resource) with the clone, you need to swap the disk signatures. You can use the Windows 2000 Resource Kit utility DumpCfg to read and write disk signatures. When you run DumpCfg, you will see the disk signature of the Storage Group volume to be replaced by the clone, as shown in the Appendix in the section “Sample DumpCfg Output.” First, a temporary disk signature is written to the Storage Group volume to be replaced. Then the original disk signature is written to the Clone volume, as shown below:

```
c:\W2kResKit>DumpCfg -S<new signature in HEX format> <disk number>
```

The example below changes the signature for Disk 1 to “17AB26CD”:

```
c:\W2kResKit>DumpCfg -S17AB26CD 1
```

**Note:** Disk Manager in Figure 12 shows a different disk numbering for the Exchange Database G: volume as Disk 2 (on Node 2). But Figure 11 shows the G: volume as Disk 1 (on Node 1). This may be the case in a cluster, so precautions need to be taken to ensure that DumpCfg is run against the correct disk.

As shown below in Figure 14, the Disk Signature for the Cluster Disk Resource is recorded in the Registry, so it is advisable to make note of the original signatures when the cluster resources are first created.

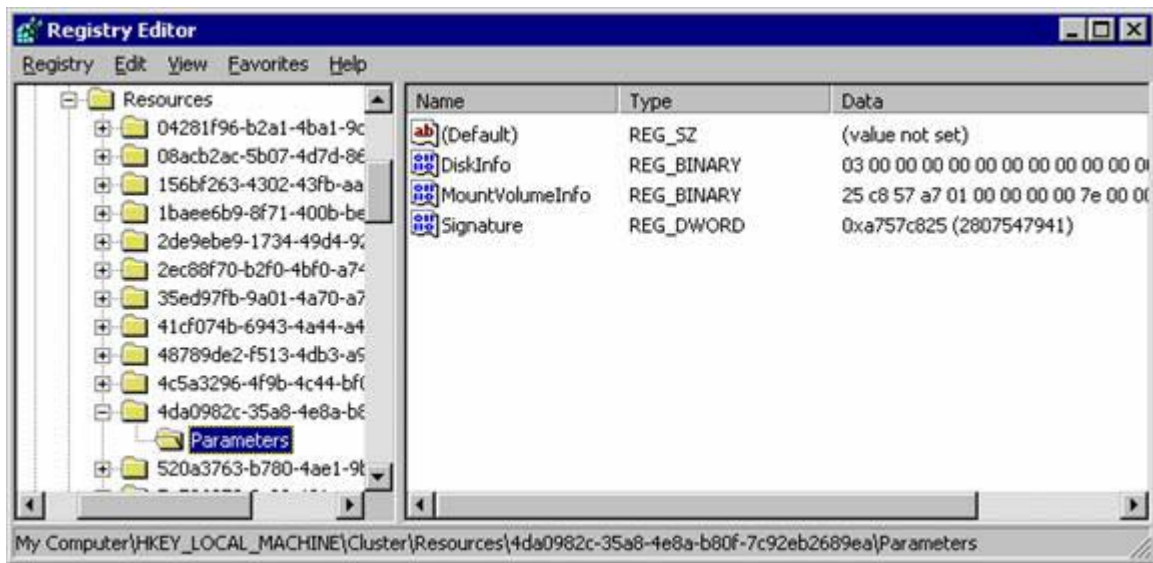


Figure 14 Location of the Cluster Disk Resource Signature in the Registry.

### ***\_CLUSTER\_NETWORK\_NAME\_ Environment Variable***

Another difference that you may encounter in a cluster environment is that the CLUSTER\_NETWORK\_NAME environment variable needs to be set to enable hard recovery.

You do not need to set this variable for a soft recovery when you recover from a disaster using a clone. However, it is recommended that you set this environment variable on each cluster node in System properties to ensure that recovery can be run when needed. (Do not wait for the time when recovery is attempted: the lack of the environment variable may prevent recovery.) See Microsoft Knowledge Base articles Q185942 and Q280770 for more information. The format for setting the variable is as follows:

```
SET _CLUSTER_NETWORK_NAME_="Virtual Server Name"
```

## **8.5 Sample Output from ESEFILE and ESEUTIL**

```
I:\Exchsrvr\sg3>eseffile /S SG3MBX1.edb
```

Microsoft(R) Exchange Server(TM) Database Utilities

Copyright (C) Microsoft Corporation 1999-2000. All Rights Reserved.

Checksumming

```

      0   10   20   30   40   50   60   70   80   90  100
      |---|---|---|---|---|---|---|---|---|---|
      .....

```

3341634 pages seen

0 bad checksums

39597 uninitialized pages

0 wrong page numbers

eseutil completes successfully after 558 seconds

**I:\Exchsrvr\sg3>eseutil /mh SG3MBX1.edb**

Microsoft(R) Exchange Server(TM) Database Utilities

Version 6.0

Copyright (C) Microsoft Corporation 1991-2000. All Rights Reserved.

Initiating FILE DUMP mode...

Database: SG3MBX1.edb

File Type: Database

Format ulMagic: 0x89abcdef

Engine ulMagic: 0x89abcdef

Format ulVersion: 0x620,9

Engine ulVersion: 0x620,9

DB Signature: Create time:02/08/2002 17:52:22 Rand:1264475 Computer:

cbDbPage: 4096

dbtime: 26576968 (0-26576968)

State: Inconsistent

Log Required: 5526-5530 ( 02/12/2002 12:40:16 )

Streaming File: Yes

Shadowed: Yes

Last Objid: 112161

Scrub Dbtime: 0 (0-0)

Scrub Date: 00/00/1900 00:00:00

Repair Count: 0

Repair Date: 00/00/1900 00:00:00  
Last Consistent: (0x1584,627,E8) 02/09/2002 05:59:03  
Last Attach: (0x1584,631,A0) 02/11/2002 13:25:03  
Last Detach: (0x0,0,0) 00/00/1900 00:00:00  
Dbid: 1  
Log Signature: Create time:02/08/2002 17:44:54 Rand:787380 Computer:  
OS Version: (5.0.2195 SP 2)

## Previous Full Backup:

Log Gen: 0-0 (0x0-0x0)  
Mark: (0x0,0,0)  
Mark: 00/00/1900 00:00:00

## Current Incremental Backup:

Log Gen: 0-0 (0x0-0x0)  
Mark: (0x0,0,0)  
Mark: 00/00/1900 00:00:00

## Current Full Backup:

Log Gen: 0-0 (0x0-0x0)  
Mark: (0x0,0,0)  
Mark: 00/00/1900 00:00:00

## Current snapshot backup:

Log Gen: 0-0 (0x0-0x0)  
Mark: (0x0,0,0)  
Mark: 00/00/1900 00:00:00

cpgUpgrade55Format: 0  
cpgUpgradeFreePages: 0  
cpgUpgradeSpaceMapPages: 0

Operation completed successfully in 0.422 seconds.

## 8.6 Sample DumpCfg Output

**Table 3 Sample DumpCfg Output.**

After Clone Split but Before Reboot	After Reboot
<pre> C:\&gt;dumpcfg  [System Information]  Computer Name: P1600D Cluster name (DNS): P1600D.domainX.net Cluster name (NetBIOS): P1600D System Root (install directory): C:\WINNT OS: Windows 2000 Server Service Pack: Service Pack 2 Product: Windows 2000 Advanced Server is installed. Terminal Services is installed.  Failure retrieving LDM disk info.   Disk Number: 4294967295 [DISKS]  Disk Number: 0 Signature: 111B39F3  Disk Number: 1 Signature: 00000001  Disk Number: 4294967295 Error - Disk Number: -1 Disk Number: 4294967295 Error - Disk Number: -1  Disk Number: 2 Signature: C174557F  Disk Number: 3 Signature: 68970E0B </pre>	<pre> C:\&gt;dumpcfg  [System Information]  Computer Name: P1600D Cluster name (DNS): P1600D. domainX.net Cluster name (NetBIOS): P1600D System Root (install directory): C:\WINNT OS: Windows 2000 Server Service Pack: Service Pack 2 Product: Windows 2000 Advanced Server is installed. Terminal Services is installed.  [DISKS]  Disk Number: 0 Signature: 111B39F3  Disk Number: 1 Signature: 00000001  Disk Number: 2 Signature: C174557F  Disk Number: 3 Signature: 68970E0B  Disk Number: 4 Signature: D418EB04  Disk Number: 5 Signature: 055E8ED5 </pre>

