COMPAQ
White Paper

December 1999
11EF-1199A-WWEN
Prepared by:
Storage Division
Compaq Computer Corporation

Best Practices for Exchange Database Management Using Compaq SANworks© Enterprise Volume Manager

Abstract: This white paper discusses how Compaq SANworks Enterprise Volume Manager helps administrators maintain their Microsoft Exchange databases for rapid recovery and high availability.

Contents

Notice

Best Practices for Exchange Database Management Using Compaq SANworks Enterprise Volume Manager
White Paper prepared by Storage Division

First Edition (December 1999)
Document Number 11EF-1199A-WWEN

Introduction

With "windows of opportunity" for performing back ups shrinking, there is a need to offload the task of performing back ups to alternative time slots and even alternative servers that won't impact the operations of a high-availability application. Additionally, there is a need to provide quick data recovery in the event of logical data corruption or physical loss of data. EVM is a utility that can automate the process of making point-in-time copies of disk volumes (also called snapshots or clones) based on software and hardware technology of the Compaq StorageWorks RAID array.

The use of both hardware and software also enables certain administrative operations to be offloaded to another server yet occur transparently to the host and host application. The use of a Storage Area Network (SAN) and the ability to attach multiple servers to the same Fibre Channel RAID array, allows EVM to migrate volumes from one server to another. For example, you might migrate a volume from an application server to a back up server or you might keep a point-in-time back up around as a "hot standby."

Note: The importance is not placed on the process that you choose to back up your data, but how long it takes you to restore the data to enable the Exchange Server to become operable once again.

How Enterprise Volume Manager Operates

Enterprise Volume Manager (EVM) is an application running on Microsoft Windows NT server (and soon Sun Solaris) that interfaces with EVM clients and with Compaq StorageWorks Fibre Channel RAID array subsystems. It utilizes in-band communications over the Fibre Channel cables through an agent to get information about devices on the RAID array. It coalesces configuration information into useful data for the EVM application. This information is essential in determining whether or not there are sufficient resources available on the RAID array to do either snapshots or clones of units.

Other features of EVM include:

- The ability to create EVM scripts for repetitive tasks
- The ability to verify that an existing script will operate with the current configuration of the RAID array controller
- "Undo" scripts to reverse the process of creating snapshots or clones
- The ability to move (migrate) volumes from one server connected to the RAID array to another server in the SAN
- An HTML-based graphical interface with output from a Web Based Enterprise Management Server (WBEM)
- The ability to execute pre- and post- command procedures on servers in the SAN
- The ability to be called from commercial back up products before and after the back up process executes

Microsoft Exchange in the Storage Area Network

Microsoft Exchange is an ideal application for utilizing a Compaq StorageWorks RAID array in a SAN. The Information Store as well as logs and MTA files can be stored on large, separate, high-speed RAID sets. Utilizing Compaq SecurePath software and multiple Fibre Channel Arbitrated Loops or Fibre Channel Switches, Exchange can be configured with no single point of hardware failure.

Although hardware can be configured for no single point of failure, the need to do back ups still exists. With EVM, administrators augment their daily back up operations and avoid lengthy downtimes due to either multiple hardware failures (double faults) or "Logical Corruptions" such as: incorrect reference counts, incorrect ACLs, a message header without a message body, and so on.

The Fibre Channel SAN model allows Exchange Server to store data in volumes on the RAID array and use EVM software to create clones of hardware-based, mirrored (RAID 1) or mirrored striped volumes (RAID 0+1) or create controller-based snapshots of any container on the Compaq StorageWorks RAID array.

Table 1. Practices for Incorporating Offline "Point-in-Back ups"

Choose either snapshot or clone volumes to back up your Information Store (PRIV.EDB, PUB.EDB).
Schedule snapshot or clones of your data immediately after any online back up.
Disable Circular Logging and use a separate volume for log files. Create a snapshot or clone of the log files volume.
Copy files from the snapshot or clone volumes to tape.
Back up Exchange support files.
Upon restore, verify the integrity of your Exchange Server database with Microsoft utilities. (ISINTEG).
Verify your ability to recovery the entire Information Store on a Recovery Exchange Server.
Verify your ability to restore individual components to a Recovery Exchange Server.

Description of Best Practices

Snapshot or Clone

The choice of doing snapshot versus clone volumes to back up your Information Store (PRIV.EDB, PUB.EDB) should be made based on what type of RAID container lives on the RAID array and how long you want the snapshot or clone to persist. In general, snapshot volumes can be created from any type of RAID container, are short-lived, and depend on the availability of the source volume being available at all times. In contrast, clone volumes begin as 3-member mirror sets, are independent of the source volume when created, and are independent of the availability of the source volume. (See table in Appendix B for a comparison of clone versus snapshot.)

Scheduling Snapshots or Clones

Based on the ability to check the logical integrity of the Exchange Server and truncate the log files after doing "online" back ups, the ideal time to make the snapshot or clone set is immediately after the back up completes. This means you will have the best possible point-in-time back up of the Information Store that you could possibly make. You should schedule your online back up to allow time to shut down Exchange, take the snapshot or clone, and then restart Exchange.

Back Up of Log Files

Your log files can be as critical as your database files for keeping your Exchange Server intact. Disable circular logging and use a separate volume for log files, then create a snapshot or clone of the logs volume. Additionally, continue saving all log files since the previous full back up. If your current full back up happens to be corrupt, you can restore the previous full back up and replay the log files. The simplest way to achieve this is to create a snapshot or clone of your log files volume when you create the snapshot or clone of your Information Store.

If you follow the rule that you must shut down Exchange before taking the snapshot or clone, you can clean up the log directory as well, since shutting down Exchange commits all data from the logs to the Information Store. If you've taken a snapshot or clone of your log files, you can safely purge files created before the previous full back up.

Note: Microsoft strongly recommends turning off "Circular Logging" on your Exchange Server. "Circular Logging" is not the default setting so you must set this on all Exchange Servers. In Exchange Server Administration check the settings on, "Site/Configuration Servers, then select File/Properties/Advanced Tab. Keep in mind that setting "Circular Logging" on the fly stops and restarts the corresponding service.

Maintaining Snapshot or Clone Volumes

Snapshot volumes depend on the source volume and the snapshot volume of the RAID array. However, the snapshot volume is meant to be temporary in its lifetime. Generally, you should attempt to preserve the data of the snapshot onto other media if you intend to maintain the data, say for several days or longer.

Clone volumes can remain around for extended periods of time since they do not depend on the source volume. However, clone volumes do not contain redundant disks, so a single disk failure will cause the clone to become inoperative.

It is both an administrative and financial decision whether to keep a snapshot or clone around or reuse it. However, having a separate snapshot (or clone) of the Information Store, the MTA and log files, and the key server volumes is ideal.

Back Up of Exchange Support Files

Exchange should not be running when you back up files in directories accessed by other Exchange services for Windows NT (such as directory synchronization – ([DX]), or Microsoft Mail message transfer agent ([PCMTA]). You should also back up your registry and Key Management Server (*\Exchsrvr\kmsdata*) and store these back ups in a safe, secure place.

Restoring the Exchange Server Database with Microsoft Utilities

Once you've restored your logs, Private Store, and Public Store to their original locations based on registry settings, you should start Exchange Directory service and change to the *\Exchsrvr\bin* directory. Then run the command:

Isinteg–patch

You may then start the Information Store service.

Restoring the Information Store on a Recovery Exchange Server

Practicing the full back up/restore process on a "Recovery Server" system than to discover that procedures do not work after a catastrophic situation. Recovery Server means installing Exchange Server to create a new site with the same site name and organization name as your production Exchange Server. However, you do not join the production site. You should have enough storage to restore the entire database and logs from any of your production Exchange Servers. Ideally, the production and recovery servers will be in the same SAN.

Begin preproduction testing by creating a small Information Store with LOADSIM rather than testing procedures with your production data. Microsoft provides this free utility on their website to assist you with the setup, configuration and size your messaging system. Once you have stabilized your test sample, you should create a snapshot or clone to test on actual production data.

For more information about the LOADSIM utility, refer to:

http://www.microsoft.com/exchange/55/downloads/LoadSim.htm

Verify the ability to restore individual components to a Recovery Exchange Server.

In many cases, the entire Information Store does not have to be restored. Instead, you may be able to recover individual components, for example, a single mailbox or a single item to a Recovery Server. You should know how to do both of these operations.

To restore a single mailbox, Microsoft recommends creating a PST file and adding it to the user's profile where they can recover whatever they need. Run the SCANPST utility against the restored PST file.

For single item recovery, set a period for deleted items on either the Private Information Store or the Individual Mailbox properties for each user. This will allow you to recover deleted items even after the user has deleted them from their mailbox. If the item is beyond the retention period, you will need to restore their entire mailbox.

Note: Unless you also perform online back ups, you should clear the "Don't permanently delete items until the store has been backed up" check box on the General tab of the Private Information Store Properties dialog box.

The reason for this is that offline back ups (back ups from a snapshot or clone) will not update the Information Store. If this box is checked, the Deleted Item Recovery will keep deleted messages around indefinitely. A typical retention time is five – seven days.

If you have been copying the Information Store and log files from the snapshot or clone to tape, you should verify that the restore operation to the Recovery Server works with files from tape.

When to Perform Snapshots or Clones with Exchange

When working with snapshots or clones on the RAID array, you must take into account a limitation with the current release of Microsoft Exchange 5.5, Enterprise Edition, Service Pack 2. In all examples that follow, EVM helps the Exchange Administrator automate the process and make it completely repeatable.

Taking Exchange Offline

In Exchange Server 5.5, there is no way to know exactly when to take the snapshot or clone unless Exchange is shut down. This shortcoming will be addressed in a subsequent version of Exchange Server (codename Platinum) with a new API being developed at Microsoft. Once a snapshot or clone has been created, it can be backed up by any software that can do file back ups.

The opportune time to shut down Exchange is when there is little or no activity on the server. If there is heavy I/O activity to the log files and Information Store, then Exchange will take a long time to shut down. Therefore, the Exchange Administrator will want to shut down Exchange properly, which entails shutting down the required services in a particular order.

After Exchange is shut down, the database is consistent since all logs have been committed to the Information Store. Once the snapshot or clone operation is complete, which should be a matter of seconds for a particular Storage Set, the Exchange Server can be restarted again, in the proper service order. (See Appendix A for sample Exchange services shut down and startup command files.)

Retaining Snapshots as a "Hot Standby"

One advantage in taking a snapshot of your Exchange database is that you have the point-in-time back up that you can revert back to without going all the way back to tape. By taking the snapshot of your data immediately after shutting down Exchange services, you also know that the logs have been committed and that the database is consistent with memory.

Some customers are indicating that they are willing to take snapshots of their Information Store several times a day. Note, however, that every time you shut down Exchange, users will generate traffic when they reauthenticate to the Exchange Server.

Snapshot Example

Suppose you begin your online back up at 12:00 A.M. and complete it at 3:00 A.M. At 8:00 A.M. the nightly back up tapes are moved offsite for safekeeping. At 7:00 A.M. you receive indication in the Error Log Utility that the Information Store is corrupt due to an unscheduled, early morning power outage that occurred during online defragmentation. This is a typical restore scenario.

1. You can retrieve your 2:00 A.M. online back up from its offsite location and restore dir.edb, priv.edb, and pub.edb. A 91-gigabyte Information Store will take a minimum of 3.5 hours to restore to disk.
2. Delete the edb.chk file and restart the Exchange Directory service (Net Start MSExchangeDS), which will replay the log files.

Instead of doing online back ups to tape, you decide to do offline file back ups of a snapshot created by the EVM software. At 12:00 A.M. you shut down Exchange services. As soon as Exchange is shut down, your EVM script creates a snapshot of the Information Store. At 12:03 A.M. you restart Exchange services and begin an NT back up of the snapshot volume on your back up server. At 7:00 A.M. you receive indication from the Error Log Utility that the Information Store is corrupt. This is another restore scenario.

1. The snapshot of the Information Store is on the RAID array controller and is available (that was not deleted or reused). Use a EVM script or the EVM GUI to migrate the snapshot volume back to the Exchange Production server. This eliminates copying any files over the network. Instead, files are moved across the SAN back into the server.
2. Copy dir.edb, priv.edb and pub.edb from the snapshot volume via a disk-to-disk copy back to the Information Store.
3. Delete the edb.chk file, restart the Directory Service to replay the log files. A 91-gigabyte information Store can be restored in as little as 10 to 15 minutes.

In either example, you should verify the integrity of the restore operation by changing to the *\Exchsrvr\Bin* directory and executing the command:

isinteg–patch

The ISINTEG utility will happen fairly quickly and is required to fix up any GUIDs in the database. If you choose to run ESEUTIL to check the database integrity after the restore operation, expect to add an additional 10 to 12 hours to the process based on Microsoft estimates.

Three-Member Mirror Sets Example

Instead of using snapshot volumes, say you decide to use controller-based striped-mirror sets (RAID 0+1) with normal operation being 3-member mirror sets. The advantage of 3-member sets is that the "split" operation to make the clone happens immediately. (See Two-Member Mirror Sets next section.)

Once you perform the clone operation, you are left with a full read-write copy of your data. After the clone set has been backed up to tape, you have two choices:

1. Select another set of drives to renormalize the mirror sets back to 3-members and leave the "hot standby" volume available but not online. Each night, you toggle between the two sets of hot standby volumes.
2. Rejoin the clone set members back to the source volume, that is no hot standby.

If you chose option 1, hot standby, you can have your Exchange Server online as quickly as you can issue commands to the controller and operating system.

Using the logical data corruption scenario above, you can make a clone volume the source of a re-normalization process on the Array Controller. Once renormalization has begun, you can set the clone online to the host and immediately replay the log files and begin verification. When the verification is complete with ESEUTIL and ISINTEG, your Exchange Server service can start. This operation will happen as quickly as the logs will replay.

Two-Member Mirror Sets Example

Instead of snapshot volumes, you instead decide to use controller-based striped-mirror sets (RAID 0+1) with the normal operation being 2-member mirror sets. The difference is that before you begin your back up process, you must begin a renormalization process of the striped-mirror set up to 3-member sets. This process will take about one hour on a 91-gigabyte volume and you can complete this operation with Exchange Server running and not cause an interruption in the service. Once the renormalization is complete, you begin the clone splitting process.

The advantage to this solution is that you do not have to switch back and forth between two different sets of devices for the same volume. The disadvantage is waiting for the normalization process before beginning the back up and clone splitting operation.

Conclusion

Providing highly reliable hardware such as the Compaq StorageWorks Fibre Channel RAID array with Compaq SecurePath software and redundant paths, hubs, and switches to the storage can greatly enhance Exchange Server availability. However, there is no way to plan for the potential of logical data corruption on the Exchange Server unless you have a solid back up plan.

The decision really depends on the following items:

- How much time can my business afford for the Exchange Server be down when trying to restore data?
- How much does it cost my business to be down during the restore time?
- How much data is stored in the Information Store, MTA directory, and log files directory?
- How much extra disk space do you need to support the snapshot or clone operation for Enterprise Volume Manager to avoid costly downtime?

Exchange Administrators can now choose to create either a snapshot or clone of their data and offload the back up operation on another server connected to the SAN. This is the “offline” approach to doing Exchange Server Back ups.

Alternatively, Exchange Administrators can continue to do “online” back ups and supplement the Exchange Server availability by retaining data from the creation of a snapshot or clone on the RAID array. If there is a data corruption and they need to completely restore their data to disk, the administrator will be able to get up and running again without having to go to tape.

Whether you choose the “online” or the “offline” method of doing back ups, you should use EVM to create the snapshot or clone of the Information Store. This point-in-time back up plus log files, plus the Microsoft provided data integrity tools, provides a complete way of preserving and restoring the integrity of Microsoft Exchange Server.

Appendix A: Sample Shut down, Back Up, and Start-up Scripts for Offline Back Up of Exchange Server 5.5

```
REM Batch File to Stop Microsoft Exchange Services
REM Note: You may not be running all of these services on
REM your server
REM
Echo Stopping Services...
Net Stop MSExchangeMSMI
Net Stop MSExchangePCMTA
Net Stop MSExchangeFB
Net Stop MSExchangeDX
Net Stop MSExchangeMTA
Net Stop MSExchangeIMC
Net Stop MSExchangeIS
Net Stop MSExchangeDS
Net Stop “PC MTA – HUB”
Net Stop MSExchangeSA
Echo Done ...
```

Snapshot or Clone Operation on Drives D:, E:, and F:

```
REM Sample Back up command using NTBACK UP
REM D is the Private Store, E is the Public Store and MTA
REM and F is the logs disk.
Ntback up BACK UP d:\ e:\ f:\ /v /d “File Based Back up” /b /l
c:\winnt\back up.log
REM Batch File to Start Microsoft Exchange Services
REM Note: You may not be running all of these services on
REM your server
REM
Echo Starting Services...
Net Start MSExchangeMSMI
Net Start MSExchangePCMTA
Net Start MSExchangeFB
```

```
Net Start MSExchangeDX
Net Start MSExchangeMTA
Net Start MSExchangeIMC
Net Start MSExchangeIS
Net Start MSExchangeDS
Net Start "PC MTA – HUB"
Net Start MSExchangeSA
Echo Done ...
```

Appendix B: Comparison of Snapshot Versus Clone Volumes

Table 2. Snapshot Versus Volumes

	Snapshot	Clone
Persistence	Short Lived–depends on availability of source volume at all times. Snapshot is lost if cache is destroyed.	Long Lived–clone is available even if source volume is deleted. Clone volume is independent of cache.
Container Type	Any container–RAID 5, RAID 0, RAID 1, RAID 0+1, JBOD	Limited Container–RAID 1 and RAID 0+1
Added Disk Overhead	Equal in size to available user volume capacity. Redundant source volume not necessarily redundant snapshot.	Equal in size to available user volume capacity. Double-redundant source volume prior to clone operation (3-member mirror sets)
Initial Construct Time	Instantaneously.	Depends on size of source volume (generally 40 minutes to 1 hour.)
Time to Split or Snap When Container is Always Available	Instantaneously.	Instantaneously–Start with 3-member mirror sets. Recreate 3-member volume in 40 minutes to 1 hour.
Reconstruct Source Volume	Must copy snapshot to another volume to revert to source volume. Unit is available after copy is complete.	Reverse mirror operation. Clone is source of copy. Unit is available immediately during re-mirroring operation.
Impact on Performance	Source volume and snapshot volume are both accessed during read and write I/O.	Source volume is accessed independently of clone volume for read and write I/O.
Required Topology	Fabric or Data Replication Manager (DRM) configuration on Target controller only.	All topologies including DRM on Initiator or Target controller.

Appendix C: Resources

Best Practices for Exchange Database Management–

http://technet.microsoft.com/cdonline/Content/Complete/srvnetwk/exch/technote/edbwp.htm

MS Exchange Disaster Recovery Part 1–

http://technet.microsoft.com/cdonline/content/complete/srvnetwk/exch/technote/edrv3p1.htm

MS Exchange Disaster Recovery Part 2–

http://technet.microsoft.com/cdonline/content/complete/srvnetwk/exch/technote/edrv3p2.htm