Technical report:

Microsoft Exchange 2007 and IBM System Storage N series

Disaster recovery solution

Document NS3584-0

February 3, 2008

Table of contents

Abstract

The purpose of this technical report is to provide a disaster-recovery scenario for Microsoft Exchange Server 2007—with cluster continuous replication setup—using an IBM N series storage solution that is designed to achieve multiple levels of RPO and RTO objectives.

Introduction

Today's business requirements, such as high availability (HA), business continuity (BC), and disaster recovery (DR), are more extensive than ever before. As Microsoft® Exchange has become a mission-critical application in recent years, Microsoft Exchange Server downtime can cost companies millions of dollars per year. IT organizations work hard to eliminate or lessen the impact of both planned and unplanned downtime through implementation of high-availability strategies and DR solutions.

Microsoft Exchange Server 2007 has matured to a new level, providing operational efficiency, delivering a messaging system that addresses performance, security, availability and cost. Microsoft Exchange Server 2007, with its new server roles and the new command-line interface, provides improved manageability and simplify maintenance. The built-in features such as local continuous replication (LCR) and cluster continuous replication (CCR) provide HA, quick recovery, and resiliency for Exchange 2007 mailbox servers. Together, IBM® System Storage™ N series and Microsoft Exchange Server increase productivity and keep information close to hand, flexible enough to meet your organization's administrative model.

This technical report delivers an overview of a DR model for Microsoft Exchange 2007 CCR using an IBM N series storage solution.

Purpose and scope

This technical report includes a disaster-recovery scenario for Microsoft Exchange Server 2007 CCR setup with an N series solution that achieves many levels of recovery point and time objectives. The scope of the solution provided is limited to the following:

- An operational DR solution for migrating Clustered Mailbox Server (CMS) onto a standby CCR cluster in a secondary location. In this case, the new CCR cluster will be in the recovery site in the DR site.
- A fully implemented DR solution for Microsoft Exchange Server 2007 based on an IBM System Storage N series solution using SnapMirror® and ReplicatorX™ (from NetApp), for replication of the production data to the DR site, and SnapManager® for Exchange (SME) and SnapDrive®, for backup and recovery.

SME restore options that are covered in this document are limited to an up-to-the-minute restore using volume IBM System Storage N series with SnapRestore® to recover database and transaction log volumes in the secondary location.

This technical report will not cover the following:

- Setup of Exchange Server recovery environment, including Microsoft Windows® Server 2003 cluster. For details on recovering Exchange 2007 CMS on a standby cluster in a DR location, see the Microsoft TechNet article: How to recover Exchange CMS on a standby cluster.
- Synchronous SnapMirror and Semi-Synchronous SnapMirror.
- Failback to the primary site.

Intended audience

This technical report is intended for information technology professionals and storage professionals responsible for corporate Exchange messaging infrastructure management. For methods and procedures mentioned in this technical report it is assumed that the reader has working knowledge of the following:

- Exchange 2007 architecture
- Exchange storage architecture and administration
- Service-level expertise of Microsoft Exchange recovery operations.

Working knowledge on IBM N series solutions, including the following:

- IBM System Storage N series with Data ONTAP®
- SnapDrive for Windows
- SnapManager for Exchange backup and restore procedures
- SnapMirror
- ReplicatorX.

Business continuity and high-availability planning

Business continuance (referred to as business continuity) describes the process and procedures an organization puts in place to ensure that essential functions can continue during and after a disaster. Continuity planning seeks to prevent disruption of mission-critical services and to reinstate full functioning, quickly and efficiently. HA is a system-design protocol and associated implementation that ensures a certain degree of operational continuance for a given measurement period.

BC and HA are not a specific technology and should integrate a variety of strategies and technologies to address all potential causes of outage, balancing cost vs. acceptable risk resulting into a resilient infrastructure. As a first step in achieving business continuity, high-availability planning is deciding which of the organization’s functions are essential to be available and operational during a crisis. Once the crucial/mission-critical components are identified, it is essential to identify your recovery point and recovery time objectives (RPO and RTO, respectively) for the identified crucial/mission-critical apportioning in terms of cost and acceptable risk. To appropriately architect a DR solution, one must be familiar with the following terms.

Availability

Generally, a degree to which a system, subsystem, service, or equipment is in an operable state for a portion of time in a functional condition. It refers to the ability of the user community to access the system.

Disaster recovery

A process of regaining access to the data, hardware, and software needed to resume critical business operations after a disaster. A disaster-recovery plan should also include methods or plans to copy required mission-critical data to a recovery site to regain access to mission-critical data after a disaster.

High availability

A system design protocol and associated implementation that ensures a certain absolute degree of operational continuity of a system, service, or equipment during a given measurement period. High-availability planning should include strategies to prevent single points of failure that could potentially disrupt the availability of mission-critical business operations.

Recovery point objective

The RPO describes a point in time to which data must be restored/recovered in order to be acceptable to the organization's process supported by the data.

Recovery time objective

The RTO is the frontier of time and service level within which service availability must be accomplished to avoid undesirable consequences associated with a break in continuity of a service/process.

Service level agreement

A service level agreement (SLA) is a formal negotiated agreement between a service provider and a user (typically customers), specifying the levels of availability, serviceability, performance, and operation of a system, service, or application.

Disaster-recovery model for Exchange Server 2007

When architecting a disaster-recovery solution for Microsoft Exchange Server 2007, it is important to review your current SLA to derive RTO/RPO objectives. We'll discuss multiple IBM N series components that were used to achieve two levels of RTO/RPO targets below.

Technology components

N series SnapManager for Exchange has achieved the "Certified Windows" logo for Windows Server 2003 for Microsoft Exchange Server backup and recovery. SME tightly integrates with Microsoft Exchange for consistent online backup in Microsoft Exchange environments while leveraging IBM System Storage N series with Snapshot™ and SnapMirror technologies. These products are critical in protecting Exchange Server data, allowing administrators to quickly back up and mirror Exchange data.

N series SnapDrive for Windows is an enterprise-class storage and data management solution for Microsoft Windows Server environments. SnapDrive enables storage and system administrators to quickly and easily manage, map, and migrate data.

N series SnapMirror: SnapMirror delivers the DR and data replication solution that today's global enterprises need. By replicating data at high speeds over LAN and WAN, SnapMirror provides the highest possible data availability and fastest recovery for mission-critical applications.

ReplicatorX: ReplicatorX is an enterprise-class solution that near synchronously replicates block data over any distance, across heterogeneous infrastructures, without operational disruption. ReplicatorX can be used for DR, data migration, and development/test or other cloning purposes.

Cluster continuous replication (CCR): CCR is a high-availability feature of Microsoft Exchange Server 2007 that combines the asynchronous log shipping and replay technology built into it — with failover management features provided by Microsoft Windows Cluster Service. CCR is limited to a two-node, active-passive cluster using the majority node-set quorum removing the shared storage-technology barrier, providing no single point of failure. Transaction logs on the active node are copied to the passive node on closing the logs and replayed on the passive node to maintain an available copy of the database.

Transport dumpster: The Transport Dumpster is an essential component for CCR and is a feature built into the Hub Transport (HT) server role. The Transport Dumpster helps recover the lost data that occurs during an automatic recovery. The Transport Dumpster takes advantage of the redundancy in the environment to reclaim the lost data from the dumpster queue maintained by the HT server.

Disaster-recovery model: objective

The primary objective of this DR model is to achieve highest degree of operational continuance at the primary site with no single point of failure and to have a recovery site and replicate the production Exchange Server data for recovery in case of a disaster. Two scenarios with multiple N series components were tested to achieve two different levels of RTO/RPO objectives outlined below.

Business case 1 (overview)

To meet a five-minute RPO and a 30-minute RTO, SME backups were scheduled and replicated to the DR site every four hours, and SnapDrive rolling snapshot copies of the transaction log volume were replicated to the DR site every 30 minutes using SnapMirror. The HT server was SAN booted and the boot LUN of the HT server was replicated every five minutes.

Business case 2 (overview)

To meet a near zero-minute RPO and a 35-minute RTO, SME backups were scheduled and replicated to the DR site every four hours, and SnapDrive rolling snapshot copies of the transaction log volume were replicated to the DR site every 30 minutes using SnapMirror. The HT server system partition was replicated near synchronously to the DR site using ReplicatorX.

Architecture

This architecture model provides HA and DR for Exchange Server 2007. It contains a primary site for production and a DR site for recovery of Exchange Server 2007. A CCR cluster in the primary site hosting 1,500 users deployed on an N series storage cluster to provide resiliency on the server hardware, application, and storage levels. A standby Windows MNS cluster deployed with Exchange passive mailbox server roles were deployed in the DR site with the N series storage cluster.

The following sections will discuss the two business case scenarios outlined earlier with implementation details, recovery methodologies, and timelines.

Disaster-recovery scenario: Exchange 2007 case study

In this section, we'll demonstrate two business cases for Microsoft Exchange Server 2007 taking advantage of the high-availability features like CCR and Transport Dumpster (TD) along with IBM N series hardware and software solutions to build a resilient infrastructure in the primary site and a DR site for recovery of Microsoft Exchange Server 2007 in case of a disaster with different RPO/RTO objectives.

Business case one: 5-minute RPO/30-minute RTO target

The following section will demonstrate a DR model for Microsoft Exchange Server 2007 using N series solutions with the architecture discussed earlier to provide the highest degree of operational continuance in the primary site and replicate the Exchange Server data to the DR site for recovery in case of a disaster with a 5-minute RPO and a 30-minute RTO objective.

The following diagram shows the basic architecture used in this scenario:

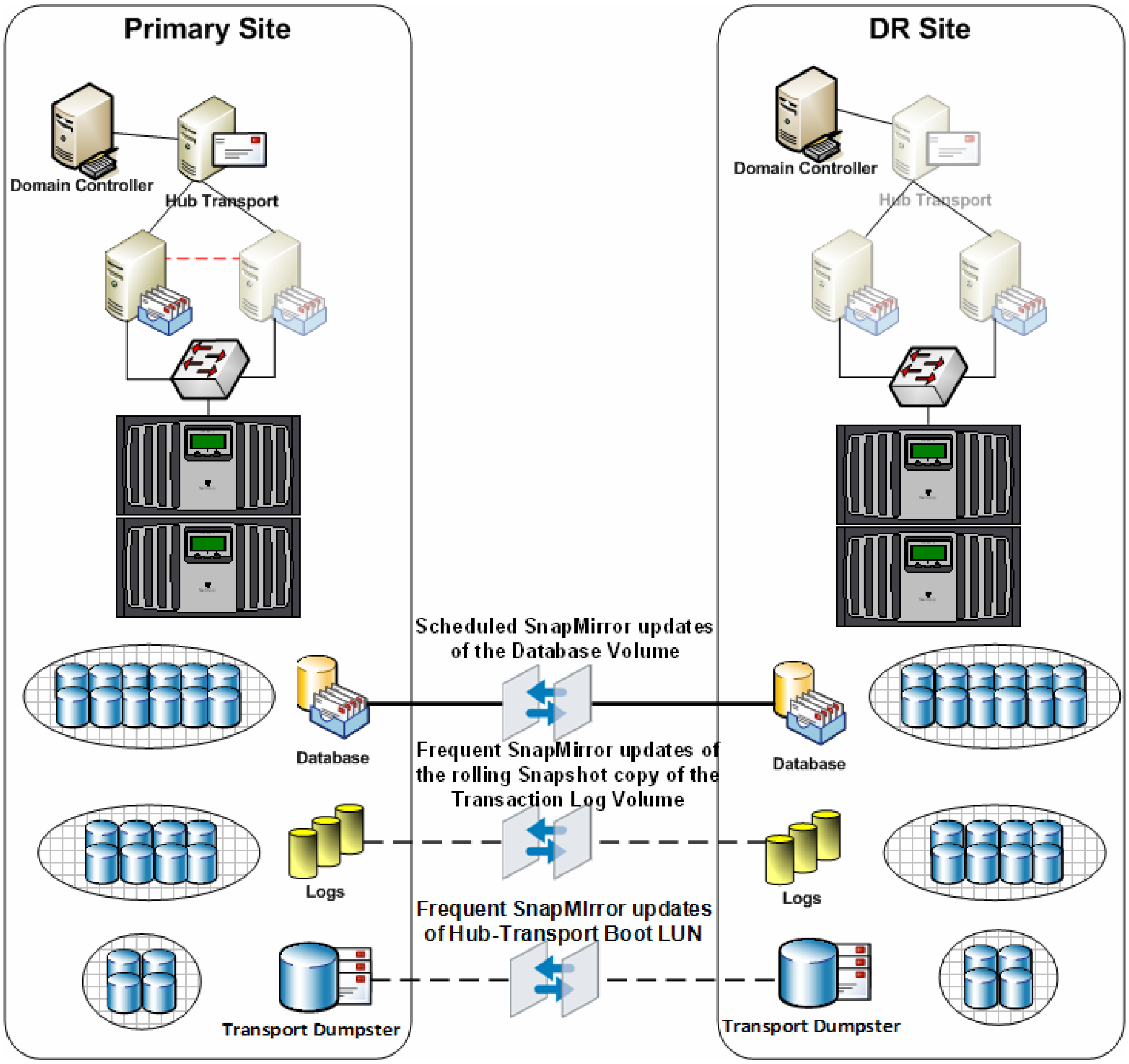

Primary Site
Domain Controller
Hub Transport
Database
Logs
Transport Dumpster
DR Site
Domain Controller
Hub Transport
Database
Logs
Transport Dumpster
Scheduled SnapMirror updates of the Database Volume
Frequent SnapMirror updates of the rolling Snapshot copy of the Transaction Log Volume
Frequent SnapMIrror updates of Hub-Transport Boot LUN

Figure 1) Business case one architecture.

Implementation details

As per the architecture discussed earlier, the following sections will provide a detailed overview of the implementation of this scenario in the primary site and the DR site.

Primary site

The primary objective of this setup is having the primary production site operational with no single-points of failure. Microsoft Exchange Server is deployed in a resilient fashion to provide resiliency in server hardware, application, and storage level.

- A Windows 2003 Active Directory forest was built using one domain controller (raised to 2000 or 2003 forest functional level).
- Two Windows Server 2003 SP1 enterprise edition with Exchange Server 2007 installed with one active mailbox role and one passive mailbox role.

- One Windows Server 2003 with Exchange Server 2007 HT (Hub Transport) server role also acting as a file share witness for the CCR cluster.
- Exchange Server databases hosted on IBM System Storage N series N7800 active-active cluster.

The following table provides a quick snapshot of the business problems and how they are addressed on the primary site providing a resilient architecture.

Business problem	Addressed?	How?	Description
Single point of failure	✓	CCR + N series storage	CCR addressing server resiliency and N series storage cluster addressing resiliency on the storage level providing no single point of failure on the application, server hardware, and storage
Disaster recovery	✓	SDW and N series SnapMirror	Replicating the database and log files to the DR site using Asynchronous SnapMirror replication, SME to provide faster backups and rapid restores
Business continuity	✓	Exchange CCR + N series storage	Exchange CCR hosted on N series storage cluster provides no single point of failure in case of a server hardware or application or a storage failure
Storage resiliency	✓	N series storage cluster	N series storage cluster provides no single point of failure on the storage level
Fast backup/recovery	✓	SME 4.0	SME as a simple SAN application integrates well with Exchange HA features providing faster backups and rapid restores
Five-minute RPO	✓	SnapMirror	Schedule SME backups every four hours and SnapDrive rolling snapshot updates using SnapMirror every 30 minutes and SnapMirror update of the HT server not LUN every five minutes
Less RTO	✓	SME/SDW and SnapMirror	Volume SnapRestore providing an instantaneous restore

Table 1) Business case one solution benefits.

HT (Hub Transport) server and Transport Dumpster (TD)

When using CCR in your environment an important step is to enable the Transport Dumpster on the HT server. Since the TD cannot be clustered or replicated at this time, we chose to SAN boot the HT server so that the HT server boot LUN can be replicated to the remote site for TD recovery in case of a disaster for the passive node to reclaim lost e-mails.

Transport database configuration

When deploying a CCR cluster, the TD configurations should be taken into consideration. The storage capacity of the HT server should contain enough capacity to store long enough for all storage groups in its site, so that messages can be recovered at the passive CCR node in the event of an unscheduled outage.

There are two important settings on the transport server that control how long a message remains in the TD:

```
MaxDumpsterSizerPerStorageGroup
MaxDumpsterTime
```

Note: By default MaxDumpsterSizerPerStorageGroup is set to 18. To size the TD properly, Microsoft recommends that you configure the MaxDumpsterSizerPerStorageGroup to 1.5 times of the max message size.

In this scenario we configured the MaxDumpsterSizerPerStorageGroup to 25MB and the max dumpster time to seven days (default).

Data replication

To achieve aggressive RPO targets for Microsoft Exchange, scheduled SnapMirror updates of the Exchange database volume and the HT boot LUN with the dumpster database and consistent SME backups replicated using SnapMirror were performed along with frequent rolling snapshot updates of the transaction log volumes.

Replication schedule

The Exchange database volumes were replicated every four hours, and SnapDrive rolling snapshot copies were updated using SnapMirror every 30 minutes and the HT boot LUN was replicated every five minutes to the DR site to make sure that at any given point of time the maximum data loss would be only five minutes. The following picture provides an overview of the replication schedules used in this architecture for a four-hour window.

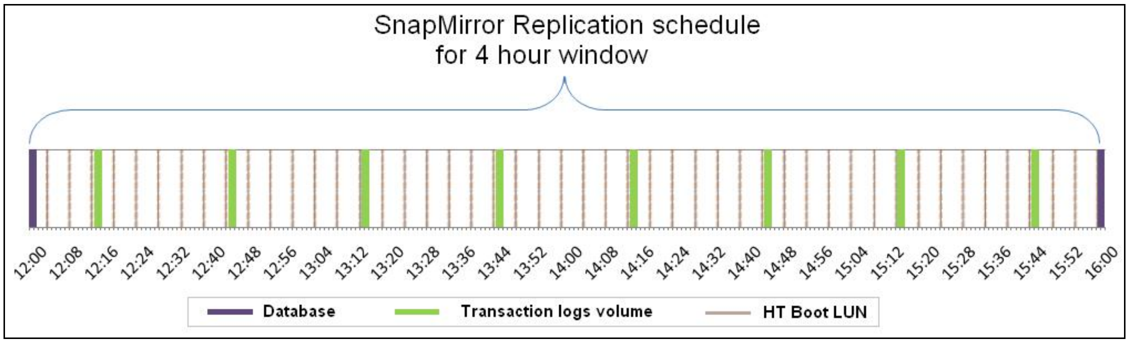


Figure 2) SnapMirror four-hour-window replication schedule.

When setting up SME backups and rolling snapshot copy schedules, it is important to take the following into consideration:

- The first snapshot copy replication must not begin until the initial backup operation is complete.
- Subsequent snapshot copy replication must not begin before the previous replication is complete. If the replication of the previous snapshot copy did not complete, is still running, then SnapMirror will wait for the replication to complete for it to start the replication of the latest snapshot.

Note: The “Snapmirror status” command can be used to monitor the status of the replication of the snapshot copies and depending on the bandwidth and data sizes, it may be necessary to change the replication schedules for the replication to complete within a reasonable amount of time.

Load generation

Load Generator (LoadGen) was used to generate a 1,500-mailbox load on the Exchange CMS in the test environment with 100 MB storage quota. After the initialization process LoadGen was run for eight hours to simulate messaging traffic. SME was configured to create snapshot copies every four hours in order to provide data need to calculate the rate of change.

DR site

As the objective of this setup is to have an operational DR site for recovery in case of a disaster**,** the following are required:

- A Windows 2003 additional domain controller
- Two-node Windows Server 2003 Enterprise Edition with SP1 MNS cluster with Exchange Server 2007 installed with passive mailbox server roles
- One server hardware identical to the HT server on the primary server to be SAN booted from the HT server Boot LUN replicated from the primary site.

Storage layout

The following storage layout was used in the test environment.

	Total capacity	Used capacity	LUNs
HT boot LUN	80 GB	56%	Boot LUN
CCR active node			
DB VOL	819 GB	54%	E:\SG1 Mailbox Store
			F:\SG2 Public Store
Logs Vol1	200 GB	49%	G:\SG1 Logs
Logs Vol2	56 GB	20%	H:\SG2 Logs
CCR passive node			
DB VOL	819 GB	54%	E:\SG1 Mailbox Store
			F:\SG2 Public Store
Logs Vol1	200 GB	49%	G:\SG1 Logs
Logs Vol2	56 GB	20%	H:\SG2 Logs

Table 2)Test storage layout.

Note: The storage layout in the DR site is an exact replica of the primary site.

Setting up SnapMirror relationship

The following section describes how to set up a SnapMirror relationship for the test environment. First create the SnapMirror destination volumes to be the DR site.

Note: These volumes have to be equal or greater than the size of the source volumes.

On the source storage controller console, use the options snapmirror.access command to specify the hostnames of the storage systems that are allowed to copy data directly from the source storage system. For example:

```
options snapmirror.access host=<destination_storage>
```

Restrict the volumes to allow SnapMirror to access them using vol restrict command:

```
Vol restrict <volume_name>
```

Initialize the SnapMirror process

From the destination storage controller console, use the SnapMirror initialize command to create an initial seed of the source on the destination and start the mirroring process.

```
Snapmirror initialize –s <src_storage_name>:<src_vol>
<dest_storage_name>:<dest_vol>
```

Note: You can use the SnapMirror Status command to check the status of the SnapMirror initialization as shown below:

```
NB-7800-1> snapmirror status

Snapmirror is on

Source               Destination          State         Lag           Status
NB-7800-1:HTPSLUN    NB-7800-2:HTRSLUN    snapmirrored  00:05:06      Idle
NB-7800-1:CCRNBDB    NB-7800-2:CCRNCDB    source        transferring  (276 MB done)
NB-7800-1:CCRNBLOGS  NB-7800-2:CCRNCLOGS  source        transferring  (56 MB done)
```

Setting up SME scheduled backups and rolling snapshot copies

SME backups can be scheduled to run from the SME console using the Windows Task Scheduler.

The following steps were taken to create an SME full backup every four hours to satisfy our DR requirements:

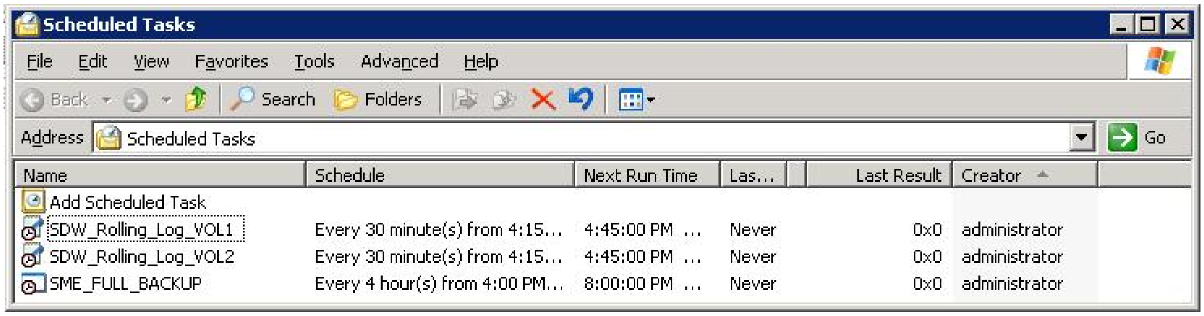
1. Open the SME console and click the Backup and Verification settings.
2. Select the storage group that you want to have backed up and replicated.
3. Ensure to check the Update SnapMirror after operation checkbox and then click the Schedule button.
4. SME will then create a new Windows task for this backup operation.

Rolling snapshot copies

The following steps outline how to create a scheduled rolling snapshot copy operation for SnapDrive using the Windows Task Scheduler.

1. Create a batch file (a file with a .bat extension) containing the following command on the Windows host on which you are scheduling the rolling snapshot copies:
   `sdcli snap update_mirror -D` *`DriveLetterList`*
2. *DriveLetterList* is a list of space-separated drive letters that contain the transaction logs.
   Example: `sdcli snap update_mirror -D g h i`
3. Select Start Menu > Settings > Control Panel > Scheduled Tasks.
4. Double-click Add Scheduled Task.
5. In the Scheduled Task Wizard, click Browse, and navigate to the folder where the batch (.bat) file you created in Step 1 is located and select it.
6. After the following panel appears, select from the list of frequencies, taking into consideration your RPO objectives, then click Next.
7. For an RPO of 33 minutes, the schedule is set to every 30 minutes.
8. Next, enter a start time and complete the detailed frequency parameters. The option details displayed on this panel vary depending on the snapshot copy frequency you picked in the previous panel.
9. Then, type the user name and password for the scheduled task, then click Next.
10. It is best to use the same SME user account and password for this task as it will have the correct permissions to execute the task.

Once you have completed the above tasks, you should have two Windows Task Scheduler jobs created and ready to execute, as seen below:



Name	Schedule	Next Run Time	Las...	Last Result	Creator
Add Scheduled Task					
SDW_Rolling_Log_VOL1	Every 30 minute(s) from 4:15...	4:45:00 PM ...	Never	0x0	administrator
SDW_Rolling_Log_VOL2	Every 30 minute(s) from 4:15...	4:45:00 PM ...	Never	0x0	administrator
SME_FULL_BACKUP	Every 4 hour(s) from 4:00 PM...	8:00:00 PM ...	Never	0x0	administrator

Figure 3) Windows Task Scheduler jobs.

Note: SME and SnapDrive do not operate concurrent operations, so it is important to make sure that the two operations do not start at the same time. If the two operations are scheduled to run at the same time, the first operation that this processed will start and the other operation will fail.

Replication of HT (Hub Transport) boot LUN

The following steps outline how to create a SnapMirror update of the HT server boot LUN using the snamirror.conf file.

We used the `Snapmirror.conf` file to schedule the SnapMirror updates of the HT server boot LUN to the DR site.

Edit the snapmirror.conf file on the destination storage using:
`wrfile /etc/snapmirror.conf` and click enter
Type `<src_storage>:<vol_name> <dest_storage>:<vol_name> - <Minute> <Hour>  <week of the month> <day of the week>`

Example: <NB-7800-1>:<HTPSLUN> <NB-7800-3>:<HTRSLUN> - 0-59/5 * * *

Note: The above example will run SnapMirror updates to the destination every five minutes.

Rate of change

Rate of change is the amount of changed data from the previous successful snapshot copy to the current snapshot copy. This information can be used to properly schedule SnapMirror times for replication of data to the DR site.

Rate of change for Exchange databases

In order to accurately calculate the rate of change for the databases volume, SME backups were scheduled and monitored for a 12-hour period. During this period the `snap delta` command can be used to estimate the rate of change between snapshot copies.

The following command can be used to display the rate of change between all snapshot copies on the database volume:

```
Snap delta CCRNCDB

Volume CCRNCDB
working...

From Snapshot                             To                    KB changed    Time         Rate
(KB/hour)
---------------                   --------------------         ----------- ------------ ---------------
exchsnap__ccr_05-18-2007_15.47.01         Active File System           9776    0d 17:55   545.544
exchsnap__ccr_05-04-2007_04.00.07 exchsnap__ccr_05-04-2007_08.00.07 2269236    0d 03:59   3725485.714
exchsnap__ccr_05-04-2007_00.00.03 exchsnap__ccr_05-04-2007_04.00.07 2636916    0d 03:59   659412.170
exchsnap__ccr_05-03-2007_20.00.03 exchsnap__ccr_05-04-2007_00.00.03 2451352    0d 03:58   616649.234
exchsnap__ccr_05-03-2007_16.00.03 exchsnap__ccr_05-03-2007_20.00.03 2173904    0d 04:01   541220.912

Summary...

From Snapshot    To                  KB changed  Time         Rate (KB/hour)
--------------- -------------------- ----------- ------------ ---------------
Snap_base        Active File System   49038424     17d 17:43   115190.059
```

The first row of the snap delta output displays the rate of change between the most recent snapshot copy and the active file system. The following rows provide the rate of change between successive snapshot copies. Each row displays the names of the two snapshot copies that are compared; the amount of data that has changed between them, the time elapsed between the two snapshot copies, and how fast the data changed between the two snapshot copies.

From snapshot copy	To snapshot copy	Size (KB)
exchsnap__ccr_05-18-2007_15.47.01	Active File System	9776
exchsnap__ccr_05-04-2007_04.00.07	exchsnap__ccr_05-04-2007_08.00.07	2269236
exchsnap__ccr_05-04-2007_00.00.03	exchsnap__ccr_05-04-2007_04.00.07	2636916
exchsnap__ccr_05-03-2007_20.00.03	exchsnap__ccr_05-04-2007_00.00.03	2451352
exchsnap__ccr_05-03-2007_16.00.03	exchsnap__ccr_05-03-2007_20.00.03	2173904

Table 3) Successive snapshot copies rate of change.

Note: If you do not specify any snapshot copies when you enter the snap delta command, the output also displays a table that summarizes the rate of change for the volume between the oldest snapshot copy and the active file system.

Rate of change for transaction logs

To calculate the amount of data generated by transaction logs, the logs volume were analyzed for four hours during the LoadGen run. Sample data was collected for four hours out of the eight-hour test run. The number of logs generated per storage group per hour was collected, and the following calculations were made.

Time	Total # of logs generated for first storage group	Total size of logs in MB
10:30 AM	1218	1218
11:00 AM	1198	1198
11:30 AM	1234	1234
12:00 PM	1247	1247
12:30 PM	1170	1170
1:00 PM	1227	1227
1:30 PM	1184	1184
2:00 PM	1187	1187

Table 4) Transaction log rate of change.

Note: When laying out Exchange data onto the storage appliance, take into careful consideration the factors outlined earlier. Things like rate of change, bandwidth available for replication, RPO/RTO for different storage groups all affect storage layout. For example, if you have executive users that require a higher RPO/RTO than normal users, its best practice to put those executive users into their own storage group, place that storage group onto its own set of LUNs, and place those LUNs into their own dedicated volumes that can be replicated to the DR site more frequently.

Recovery methodology in the event of a disaster

This section will outline the recovery methodology on the DR site in case of a disaster. The following steps must be taken prior to running recovery on the DR site.

LUN drive letters must be the same as the primary site.

SME configuration must be completed.

Volume SnapRestore

When failing over to a secondary site and restoring the whole Exchange cluster, the Volume SnapRestore method can be used as a means of decreasing the time required to complete the restore while minimizing overhead on the storage controller.

The following steps were used to recover the Exchange CMS at the DR site using Volume SnapRestore method:

Preparation

1. Issue a `SnapMirror break` of the HT boot LUN, Exchange Server DB, log, MTS/SMTP, and SnapInfo volumes.

   Example: `snapmirror break <Vol_Name>`

Recover the HT (Hub Transport) server

1. Remove any LUN mapping carried from the primary site and boot the HT server on the replicated boot LUN.

Recover the Exchange CMS

1. Perform a Volume `SnapRestore` of the Exchange Server log volume(s) with the most recently completed SnapDrive initiated snapshot copy.

   Typically this will be the rolling snapshot copy.
   If a site failover is initiated shortly after an SME backup but before the next rolling snapshot copy is initiated, there may a case in which the `SME_eloginfo` and/or `_recent` snapshot copy is more recent than the rolling snapshot copy.

2. Ensure all recovered volumes are online.

3. Clear all LUN mappings that may have been carried over from the production site.

4. Turn `SnapMirror off` on the storage appliance.
   Note: If this is not possible due to other SnapMirror relationships, remove the Exchange destination volumes from the snapmirror.conf file.

5. Connect all the LUNs on the first recovery node, which will recover and host the Exchange CMS at the DR site using SnapDrive.

6. From the recovery node verify that all the LUNs are connected.

7. To recover the Clustered Mailbox Server (CMS), run **`setup.com /recoverCMS /CMSName:<CMS_Name> /CMSIPaddress:<IP_Address>`** from the bin directory under the Exchange program files.

SnapManager configuration and restore

1. Start SME and rerun the configuration wizard.
   Note: Ensure that the Snapinfo directory matches the production site. This is especially important if a dedicated SnapInfo LUN is used.

2. Restart SME and select Restore and select “First Storage Group” for the most recent snapshot copy and click Restore.
   Note: Ensure that the up-to-the minute restore option is selected and the recover and mount database after restore option is unchecked and click Restore.

3. Repeat the previous step for all the storage groups.

4. Using Exchange Management Console mount all of the Exchange storage groups.
5. Perform an SME backup of the newly recovered Exchange environment.

Restore time

A disaster simulation was perfumed two hours 15 minutes after the last SME backup and the following metrics were observed during the recovery at the DR site to track the total recovery time and the recovery point for the Exchange CMS:

Initial steps	Time to completion
Break the SnapMirror relationship for all volumes	30 seconds
Map LUN and boot HT server	4 minutes
Perform a SnapRestore operation of all volumes	1 minute
Clear LUN mappings that may have carried from the primary site	1 minute
Connect and mount all LUNs on Exchange Server	5 minutes
Recover Exchange CMS	2 minutes
Total time	13:30 minutes

Table 5) Restore steps time to completion.

Storage group	# of transaction logs to replay	Size of the transaction log directory	Log replay time	Total SME restore time
First storage group	4863	4.74 GB	11:30 minutes	15:20 minutes
Totals	4863	4.74 GB	11:30 minutes	15:20 minutes

Table 6) Total restore time.

6. Open Exchange Management Shell and run **`Update-StorageGroupCopy identity:<DomainName>\<CMSName>`** to seed the passive node.

The total recovery time for the scenario tested was approximately 27 minutes and the recovery point objective was 15 minutes without the TD and five minutes with the TD being replicated every five minutes from the production site.

When planning RTO and RPO, it is important to know the approximate number of logs that would need to replay and the rate at which the logs replay. In this recovery scenario, it approximately took 11 minutes to replay the logs and the total restore time was 15 minutes. From this data we can calculate the logs for this test scenario were replayed at a rate of approximately 405 logs per minute or 405 MB per minute.

Business case two: Zero-minute RPO/30-minute RTO target

The following section will demonstrate a DR model for Microsoft Exchange Server 2007 with the same architecture discussed earlier to achieve a zero-minute RPO and a 35-minute RTO targets with the use of ReplicatorX.

The following diagram shows the basic architecture used in this scenario:

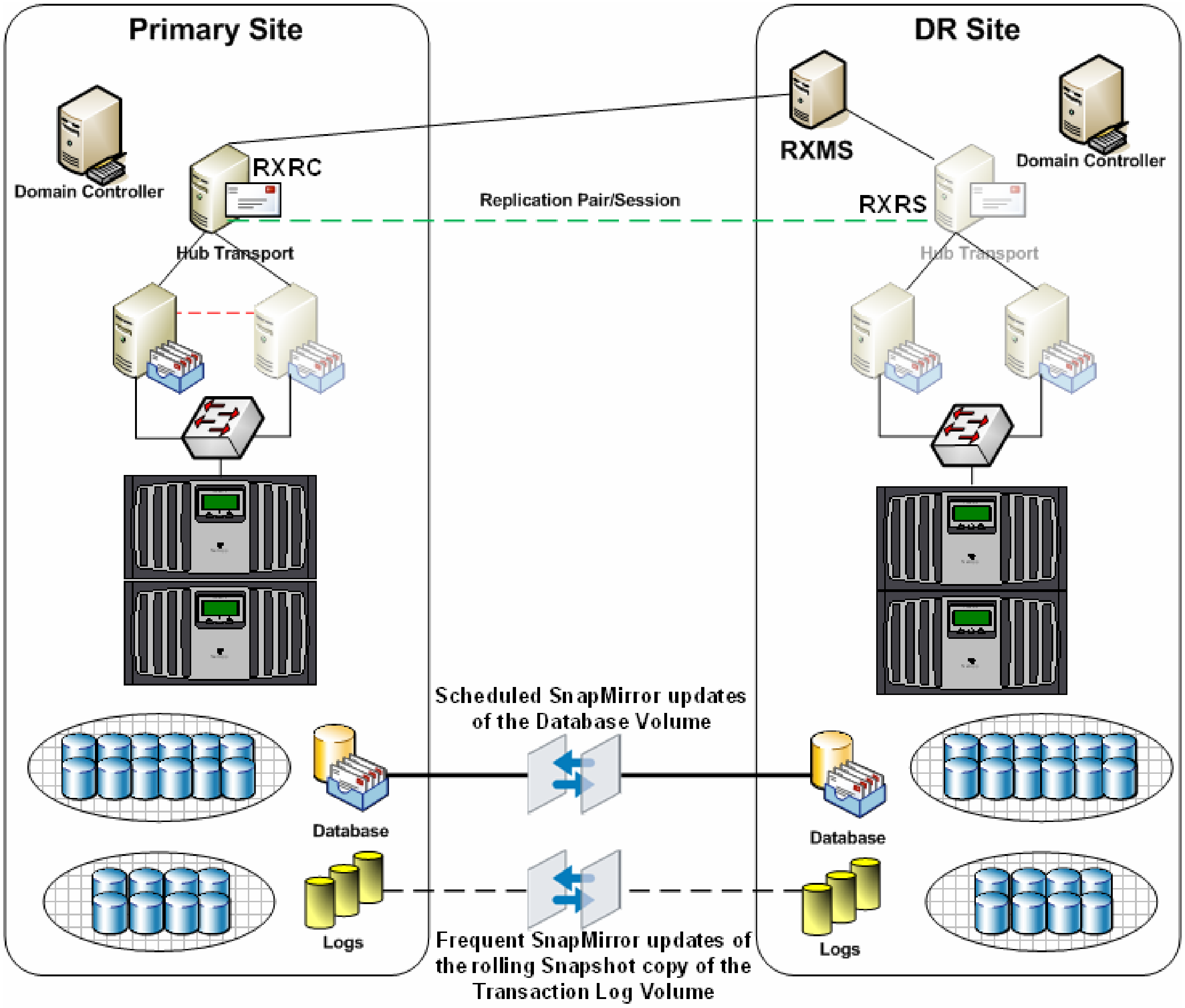


Figure 4) Business case two architecture.

Implementation details

The implementation in this business scenario is an exact replica as the previous architecture except for the HT server replication. ReplicatorX was used to replicate the HT server to the DR site continuous asynchronous replication and provide the ability of replicating the system partition of the

HT server to either IBM N series or third-party storage. The following section will describe the implementation of ReplicatorX to replicate the HT server data to the DR site.

The following table provides a quick overview of the business problems and how they are addressed on the primary site providing a resilient architecture.

Business problem	Addressed?	How?	Description
Single point of failure	✓	CCR + N series storage	CCR addressing server resiliency and N series storage cluster addressing resiliency on the storage level providing no single point of failure on the application, server hardware, and storage
Disaster recovery	✓	SDW and N series SnapMirror	Replicating the database and log files to the DR site using Asynchronous SnapMirror replication, SME to provide faster backups and rapid restores
Business continuity	✓	Exchange CCR + N series storage	Exchange CCR hosted on N series storage cluster provides no single point of failure in case of a server hardware or application or a storage failure
Storage resiliency	✓	N series storage cluster	N series storage cluster provides a No-Singe-Point of failure on the Storage level
Fast backup/recovery	✓	SME 4.0	SME as a simple SAN application integrates well with Exchange HA features providing faster backups and rapid restores
Zero RPO	✓	SnapMirror and ReplicatorX	Scheduled SnapMirror updates of the database LUN and rolling snapshot updates of the transaction logs LUNs with continuous asynchronous replication of HT server using ReplicatorX
35-minute RTO	✓	SME/SDW and SnapMirror	Volume SnapRestore providing an instantaneous restore

Table 7) Business problem solutions.

Replication of HT (Hub Transport) server using ReplicatorX

ReplicatorX release 4.0 was used in this architecture to replicate the HT server data to the DR site.

ReplicatorX Components

RXRC: ReplicatorX replication client

Installed on the each primary site server which has direct access to the source volumes.

RXRS: ReplicatorX replication server

Installed on at least one server in the secondary site which has access to the target volumes.

RXMS: ReplicatorX management server

Each secondary site must have at least one server with RXMS installed on it.

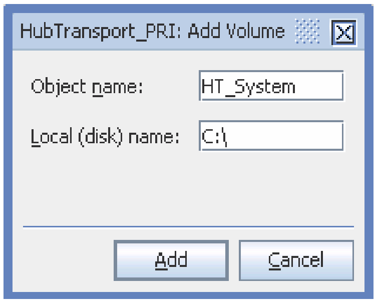
The following steps outline the how to replicate the HT server to the DR site using ReplicatorX.

Install RXRC component on the HT server in the primary site.

1. Install RXRS component on the recovery server which has access to the target volume for the HT volume.
2. Install RXMS on an additional server to manage and administer ReplicatorX.
3. Open the ReplicatorX management GUI on the RXMS server.
   a. Open Internet Explorer.
   b. Type http://<Servername>/RepX_R4_0.
4. Add the primary site and add the RXRC client to the primary site.
5. Add the DR site and the RXRS client to the DR site.

Preparing RXRC

1. Right-click the RXRC server and click Add Volume.
2. Specify a name for the volume and specify the source volume. In this case it is the C drive of the HT server.



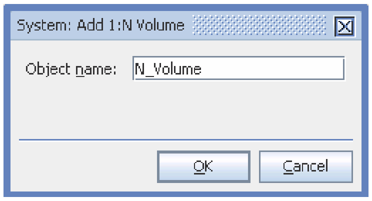
HubTransport_PRI: Add Volume

Object name: HT_System

Local (disk) name: C:\

Figure 5A) Volume name.

3. Right-click Add Volume and click Add N: Volume.



System: Add 1:N Volume

Object name: N_Volume

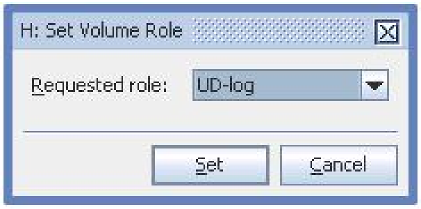
4. *Figure 5B) Volume add.*Right-click the additional volume on the RXRC server and click Set Role, and select Bitmap Role and click OK.
5. Right-click the RXRC server and click Add Bitmap File and specify the file name, size and location. (Bitmap volumes should be RAW volumes and should not be formatted with a file system.)

Note: In this scenario the bitmap volume size was 2.5 GB (approximately 3% of the source volume). It is recommended to follow proper sizing and capacity planning guidelines when sizing bitmap volumes.

Preparing RXRS

The RXRS server requires two volumes (UD-LOG and MD-LOG).

1. Right-click a volume in the RXRS server and click Set Role. Then, select UD-Log and click OK.
2. Right-click another volume in the RXRS server and click Set Role. Then, select MD-Log and click OK.



H: Set Volume Role

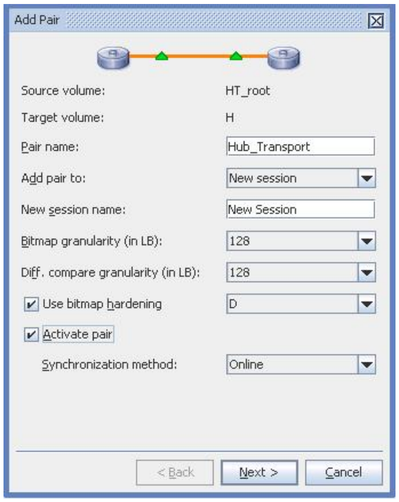
Requested role: UD-log

*Figure 5C) Volume role.***Note:** In this scenario the MD-Log volume size was 2.5 GB (approximately 3% of the source volume) and the UD Logs volume size was 14 GB (approximately 20% of the source volume). It is recommended to follow proper sizing and capacity planning guidelines when sizing bitmap volumes. (UD-Log and MD-Log should be RAW volume and should not be formatted with a file system.)

Note: Target Volume should be a Primary basic partition and should not be formatted with a file system

Creating and activating a session/pair

1. From the RXMS GUI, click Graphic-Configuration Mode, and select the volume which you want to replicate and drag It to the target volume.
   Note: The target volume should be of equal size or greater than the source volume.
2. Add Pair Dialog appears, specify the pair name and session name, check the Activate Pair check box and set the synchronization method to "**Online**" and click Next.



Add Pair

Source volume: HT_root

Target volume: H

Pair name: Hub_Transport

Add pair to: New session

New session name: New Session

Bitmap granularity (in LB): 128

Diff. compare granularity (in LB): 128

☑ Use bitmap hardening: D

☑ Activate pair

Synchronization method: Online

Figure 5D) Add pair.

3. Select the compression method to be enabled and the DPR mode to slim only and click Finish.

4. Upon clicking Finish, the initial synchronization starts.

Recovery methodology in the event of a disaster

The recovery methodology in this architecture is same as the previous business case. The following section will outline the recovery methodology of the HT server using ReplicatorX.

Recover the HT (Hub Transport) server

1. Open the RXMS Management GUI.

2. Right-click the session in the Graphic Configuration mode and click the session name and click Freeze session.

3. Once the freeze is completed there will be an ❄ icon displayed on the target volume as shown in the figure below.

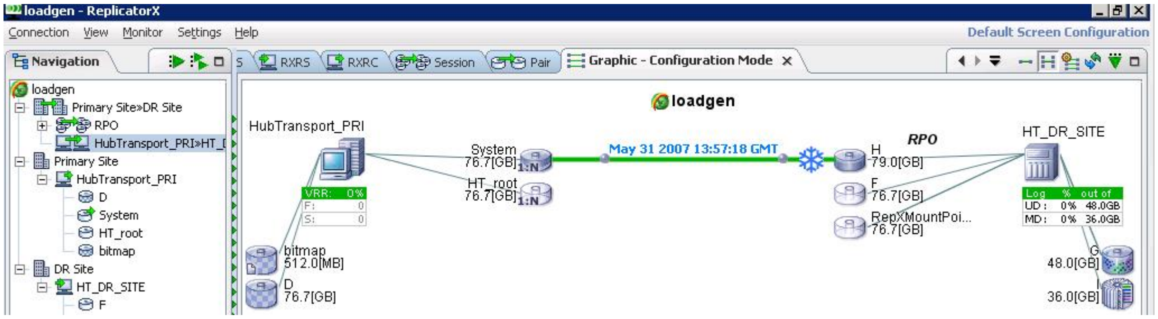


Figure 6) Freeze session.

4. Login to the RXRS server and from the command line change directory to <X:\Program Files\RepX\R4_0\RepX Replication Server\bin> and run the disk signature command to change the 0th sector of the disk to make if usable.

```
disksignature tv=d:\ fixall=1
```

Note: Refer to ReplicatorX Administration Guide for more information on the Disk Signature command like and options.

5. Reboot the server from the target volume (Which in this case is a Secondary hard drive on the hub-transport server.

Note: If the target volume resides on N series Storage, then the volume should be mapped to the DR server and booted up.

Exchange CMS (Clustered Mailbox Server) was recovered using the steps described in the previous business case.

Restore time

A disaster simulation was perfumed two hours 15 minutes after the last SME backup and the following metrics were observed during the recovery at the DR site to track the total recovery time and the recovery point for the Exchange CMS:

Initial steps	Time to completion
Break the SnapMirror relationship for all volumes	30 seconds
Freeze the ReplicatorX session	10 seconds
Run DiskSignature.exe to recover the target volume	10 seconds
Reboot the server from the target volume	5 minutes
Perform a SnapRestore of all volumes	1 minute
Clear LUN mappings that may have carried from the primary site	1 minute
Connect and mount all LUNs on Exchange Server	5 minutes
Recover Exchange CMS	2 minutes
Total Time	**14:50 Minutes**

Table 8) Restore steps time to completion (business case two).

Storage group	# of transaction logs to replay	Size of the transaction log directory	Log replay time	Total SME restore time
First storage group	4863	4.74 GB	11:30 minutes	15:20 minutes
Totals	**4863**	**4.74 GB**	**11:30 minutes**	**15:20 minutes**

Table 9) Total restore time (business case two).

The total recovery time for the scenario tested was approximately 30 minutes 10 seconds and the recovery point objective was 15 minutes without the TD and nearly zero with the TD being replicated using ReplicatorX asynchronously to the DR site.

When planning RTO and RPO, it is important to know the approximate number of logs that would need to replay and the rate at which the logs replay. In this recovery scenario, it approximately took 11 minutes to replay the logs and the total restore time was 15 minutes. From this data we can calculate the logs for this test scenario were replayed at a rate of approximately 405 logs per minute or 405 MB per minute.

Summary

Microsoft Exchange is a mission-critical application and it can cripple the operational productivity if it becomes unavailable. IBM System Storage N series has proven data protection and DR tools for Microsoft Exchange. SnapManager for Exchange backup and restore capabilities combined with SnapDrive, SnapMirror technologies and ReplicatorX provide a solid and robust solution for protecting and recovering your exchange data while meeting stringent RPO and RTO objectives based on your business requirements.

Appendix A: Best practices

When planning and sizing a DR solution for Microsoft Exchange Server environments the following best practices should be considered.

- Determine which data to be replicated and the replication methods (Synchronous or Asynchronous). This can impact the available bandwidth and RPO/RTO objectives.
- Plan volume layout to help archive RPO/RTO objectives: Separate users, business units requiring faster access and minimal data loss into separate storage groups and volumes. This adds flexibility to SnapMirror schedules and recovery objectives.
- Properly plan and size SnapMirror replication and schedules: Determine rate of change for each volume to ensure that the amount of data to be transferred by the SnapMirror process fits within desired incremental update times.
- When determining RTO objectives, ensure that the following processes are taken into account.
  - Time required for LUN recovery process.
  - Amount of time required to complete the LUN clone split process.
  - Number of transaction logs to be replayed during an up to the minute recovery Increase the frequency of rolling snapshot copies to decrease the amount of data to be replicated as well as the amount of transaction logs to be replayed during the restore process.
- When laying out your Exchange data onto the storage appliance, take into careful consideration the factors that were outlined in the above sections. Things like rate of change, bandwidth available for replication, and RPO/RTO for different storage groups all affect storage layout. For example, if you have executive users that require a higher RPO/RTO than normal users, it is best practice to put those executive users into their own storage group, place that storage group onto its own set of LUNs, and place those LUNs into their own dedicated volumes that can be replicated to the DR site more frequently.
- When scheduling the SME backups and the rolling snapshot copy backup jobs, stagger the run time so the two operations do not start at the same time. SME and SnapDrive do not support concurrent operations. If two operations do occur at the same time, the operation that is processed first will succeed; the other operation will fail.

Trademarks and special notices

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Information is provided "AS IS" without warranty of any kind.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.