White Paper

EXCHANGE 2010 DISASTER RECOVERY OPTIONS WITH CROSS-SITE EXCHANGE DAG AND EMC RECOVERPOINT

A Detailed Review

EMC GLOBAL SOLUTIONS

Abstract

This white paper provides analysis, data points and design considerations on how a Microsoft Exchange Server 2010 disaster recovery (DR) strategy can be deployed in cross-site configurations. Two options are discussed and examined—Exchange 2010 with cross-site Database Availability Group (DAG) protection, and Exchange 2010 with local site DAG protection using EMC® RecoverPoint for cross-site replication and recovery.

April 2011

Copyright © 2011 EMC Corporation. All Rights Reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

The information in this publication is provided "as is." EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

VMware, ESX, and vSphere are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions.

All trademarks used herein are the property of their respective owners.

Part Number H8104

Table of contents

Executive summary

E-mail communication is probably the most critical part of any business. When Microsoft Exchange Server 2010 is deployed it becomes the most important and crucial part of the company communication infrastructure requiring dedicated 24x7 support staff. Exchange administrators are responsible for keeping the Exchange system up and running to ensure consistent business continuity, and communications in the event of unforeseen disasters. So how can you, as an Exchange or e-mail infrastructure administrator, ensure data protection, fast recovery and high availability (HA) if unplanned outages or natural disasters occur?

To help answer these questions, the Exchange administrator needs to be familiar with the key HA functionality features available in Exchange 2010, and understand how they may impact the organization's recovery time objectives (RTO), and recovery point objectives (RPO).

EMC offers multiple business continuity and disaster recovery (DR) solutions for different applications, including Microsoft Exchange Server 2010. Depending on your organization's service level agreements (SLAs), the Exchange administrator can choose the best possible options that can help to meet or exceed RTO and RPO requirements. This white paper presents Exchange 2010 administrators with analysis, design considerations, and replication performance data points for two possible Exchange 2010 DR options:

- Exchange host-based DAG replication to a remote DR site with a cross-site DAG configuration over high latencies (up to 250 ms round-trip)
- Exchange replication to a remote DR site with EMC RecoverPoint CRR over high latencies (up to 250 ms round-trip)

Introduction

Purpose

The purpose of this white paper is to review design guidelines and performance results identified during validation of two Exchange 2010 DR options:

- Exchange 2010 mailbox resiliency with native DAG replication for local and cross-site protection
- Exchange 2010 mailbox resiliency with native DAG at local site with cross-site replication provided by EMC RecoverPoint CRR

Scope

This white paper covers the following topics:

- Exchange 2010 DAG design in a virtualized environment
- Exchange 2010 host-based DAG replication to a remote DR site over high latencies (up to 250 ms round-trip)
- Exchange 2010 replication performance to a remote DR site with RecoverPoint CRR over high latencies (up to 250 ms round-trip)
- Exchange 2010 procedures to activate database copies at the DR site in a cross-site DAG configuration
- Exchange 2010 DR procedures using database portability with RecoverPoint CRR

Audience

This white paper is intended for these audiences:

- EMC employees
- EMC partners
- Customers, including IT planners, storage architects, and Exchange administrators
- Field personnel tasked with designing and implementing a DR solution for Microsoft Exchange Server 2010 on EMC CLARiiON® storage

Considerations when deploying Exchange 2010 DR solutions

Exchange 2010 HA and site resiliency concepts

Exchange Server 2010 introduces a new concept called Database Availability Group (DAG). DAG HA and site resiliency functionality is designed to provide business continuity, should an unplanned database failure, or mailbox server failure occur in your Exchange environment. A DAG is a set of mailbox servers (up to 16) that replicate to one another, providing database-level protection from outages.

DAG uses Exchange Continuous Replication (logs shipping) and a subset of Windows failover clustering technologies to provide continuous HA and site resiliency. It replaces earlier Exchange and Windows failover clustering based technologies used for HA in Exchange 2007, such as single copy clusters (SCC), continuous cluster replication (CCR), and standby continuous replication (SCR). DAG members can reside in the same data center or they can be configured across multiple sites in cross-site, "stretched" configurations.

It is important to note that while Microsoft is relying on DAG to provide an optimal, Exchange 2010 cross-site DR solution, our findings show that RTO could be significantly increased if multiple databases need to be replicated across high latency networks with limited bandwidth. Based on the above facts, questions arise as to whether a DAG-only solution can truly fulfill all of an organization's e-mail data recovery needs. In addition, deploying cross-site DAG only provides a solution for the Exchange 2010 servers when most customers prefer a single replication solution for managing and monitoring replication availability of all other critical applications, including data residing on Windows, Linux, and UNIX systems.

Network considerations for site resiliency solutions

When designing any DR solution, it is important to understand the capacity, speed, and utilization of the link between the production and DR sites. In many cases, the link is the factor that causes a solution to fail.

When designing Exchange 2010 for cross-site replication, Exchange administrators need to be aware of the specific network requirements and conditions that may prevent successful replication. During our testing we observed that replication to a remote site can be challenging. Specifically, this may impact solutions that are leveraging DAGs in a cross-site (stretched) configuration where the network latencies are above 100 ms and bandwidth is less than 1 Gb/s. We suggest you use the Microsoft Exchange 2010 Mailbox Role Calculator to determine necessary network link requirements when planning to implement Exchange 2010 in a stretched DAG solution.

Important: Microsoft recommends that the log replication throughput metrics are identified when planning for DAG replication. The throughput metrics are dependent upon knowing the proper log generation rate per hour of the day for your environment. If this data is unknown, then the log replication throughput metrics provided by the calculator may not be accurate. Information about the Exchange 2010 Mailbox Role Calculator can be found at:

http://blogs.technet.com/b/exchange/archive/2010/01/22/3409223.aspx

For additional information about deploying Exchange 2010 in a mailbox resiliency configuration with DAG, see:

http://technet.microsoft.com/en-us/library/dd979799.aspx

EMC RecoverPoint Technology

EMC RecoverPoint supports mixed host operating system environments including Linux, Windows, Solaris, UNIX and VMware®, and provides heterogeneous bi-directional replication between storage arrays from multiple array vendors. RecoverPoint also provides local continuous data protection (CDP), remote (CRR), or concurrent local and remote (CLR) data protection with the capability to replicate one or more production LUNs. All replicas can have different protection windows and RPO policies as each replica has its own policy settings. These policies enable the customer to set different RPOs for the local and remote copy.

Unlike other replication products, RecoverPoint is appliance-based, which enables it to better support large amounts of information stored across a heterogeneous server and storage environment. RecoverPoint uses lightweight splitting technology—on the application server, in the fabric, or in the EMC VNX™ series, CLARiiON CX4™ or CX3 arrays. This technology is used to mirror a write to a RecoverPoint appliance that resides outside of the primary data path. Implementing an out-of-band approach enables RecoverPoint to deliver continuous replication without impacting an application's I/O operations. If a splitter based on VNX or CLARiiON technology is used, then SCSI and iSCSI LUNs hosted by the array can be used by RecoverPoint.

The EMC network-based approach to data protection enables certain key functionalities, such as being able to instantly recover data to any point-in-time, by leveraging a journal that stores all data changes and bookmarks that identify application-specific events. Data can be recovered as read/write without affecting the application servers or ongoing replication. RecoverPoint makes use of consistency groups (CGs), which let users define sets of volumes for which the inter-write order must be retained during replication and recovery. This ensures that data at any point in time will be fully consistent. Finally, the RecoverPoint appliance provides advanced bandwidth management, which implements a policy engine that decides how much data should be compressed before sending it across an IP or Fibre Channel (FC) link.

EMC offers a variety of innovative and advanced ways to manage your Exchange storage, reduce costs, and protect your data. Organizations implementing EMC RecoverPoint for Microsoft Exchange can expect to see the following benefits:

- Elimination or significant reduction of backup windows along with cost reductions by minimizing removable media usage
- Improvements in RPOs and RTOs with minimal application downtime during a disaster, or planned server or storage upgrades
- Use of the VSS framework when replicating Exchange ensures that the RecoverPoint images meet Microsoft requirements for Exchange integrity and recoverability
- Full read/write access to all point-in-time snapshots, which allows the user to dial back in time and perform e-mail searches, migration testing, or server

recovery without impacting the performance of the production servers, or the replication process

- Out-of-band processing for replication ensures that Exchange Server performance is not impacted by RecoverPoint replication processes
- Rapid and simple replication for backup and recovery, automatically synchronized to an alternate location, and instantly accessible for DR, or for recovery from logical corruption

More information on using EMC RecoverPoint for Exchange server recovery can be found at the RecoverPoint web page on www.EMC.com, see:

- *EMC RecoverPoint Family Overview –A Detailed Review* white paper
- *Improving Microsoft Exchange Server Recovery with EMC RecoverPoint —A Detailed Review* white paper

Items to consider before choosing the cross-site DAG DR option

Depending on customer SLAs, RPO and RTO requirements, a solution with cross-site DAG, limited bandwidth, and high latencies between sites might fall short of customer needs. Review some of the real-world scenarios detailed in Table 1 to identify some of the DAG limitations, and how RecoverPoint can help to overcome them.

Table 1. Considerations for a stretched DAG option and RecoverPoint advantages

Customer scenario	DAG limitations and considerations	RecoverPoint advantages
Customer requires daily, full backups.	Logs truncation cannot occur during the database seeding process. The full backup will fail if the seeding process takes longer than the allocated nightly backup window. RPO and RTO requirements will not be met.	Replication is performed at the block-level and does not affect the mailbox server. Backups can be performed without any problems. RPO and RTO requirements are met.
There is an interruption on the replication network between production and DR sites.	• If database being seeded, the seeding process will fail, and require a restart from the beginning. • An existing database copy will be suspended, and logs will start to accumulate until the link is re-instated. The database could be dismounted if the logs LUNs are not sized properly to accommodate these conditions.	Replication will be paused and restarted after the link is re-established. Only changed tracks are replicated which significantly improves replication times and reduces RTO. There is no negative effect to the Exchange host because replication is performed at the block-level.
Customer environment runs multiple versions of Exchange (Exchange 2010 and earlier).	Earlier Exchange versions require different replication and HA solutions. DAG only protects Exchange 2010 mailbox servers.	Replication is performed at the block-level and is not dependent on the version of Exchange. All Exchange versions can be replicated in one single solution.
Customer requires protection for Exchange public folder database.	DAG provides protection for the mailbox database only.	Replication is performed at the block level and host agnostic. RecoverPoint can replicate any mailbox data, as well as the public folder data.
There are high latencies on the replication network between primary and DR sites. Network bandwidth is limited.	DAG is not WAN-optimized, nor does it offer any bandwidth tuning or replication scheduling capabilities. Also, there may be distance limitations depending on network latency requirements.	RecoverPoint provides WAN-optimization, bandwidth throttling and replication scheduling to support both business continuity and DR strategies while minimizing network requirements and costs. RecoverPoint also provides failover and failback with built-in compression.

RecoverPoint CRR for Exchange 2010 DR

EMC RecoverPoint can provide Exchange 2010 protection for both local and remote copies. In this white paper, we will focus on providing protection for remote DR copies using RecoverPoint in Continuous Remote Replication (CRR) mode.

EMC RecoverPoint CRR is integrated into a customer's existing network infrastructure to allow the replication of multiple applications to farther distances on cross-site networks with high latencies. Specifically, for Microsoft Exchange 2010, RecoverPoint in CRR mode can be used to asynchronously replicate Exchange 2010 database copies to a remote DR site without using stretched DAG. During a site failover event the Exchange 2010 *database portability* process can be used to activate replicated copies on a stand-alone mailbox server, or another Exchange mailbox server that is a member of a different DAG at that site.

Database portability is a feature that enables the Microsoft Exchange Server 2010 mailbox database to be moved to, or mounted on any other mailbox server in the same organization. This feature can help reduce overall recovery times for various failure scenarios. For more details on the database portability process, see http://technet.microsoft.com/en-us/library/dd876873.aspx.

RecoverPoint CRR can also replicate an entire Exchange DAG, and Exchange virtual machines across the WAN to allow for the entire Exchange infrastructure to be available at the DR site in case of primary site critical failure. Replicated data can be used for backups, archiving, mailbox recovery, or other developmental reasons.

Simulated Exchange 2010 test environment profiles

To help customers identify the most effective DR solution, we have created two real world Exchange configurations in our lab. Both test environments incorporate a WAN situated between a primary, and remote DR site where a distance emulator simulates high latencies with multiple link types ranging from 1 Gb/s Ethernet to 155 Mbps OC-3.

During testing, we have injected multiple latencies (up to 250 ms) and performed the initial database replication process (seeding with DAG and "full sweep" with RecoverPoint), site failover operations to evaluate how this process may impact RPO and RTO requirements. In both tested solutions we have virtualized Exchange 2010 by using VMware vSphere™ 4.1, and leveraged Microsoft best practices and guidelines for deploying DAG in multisite configurations.

For more details on Exchange 2010 deployments for HA and site resiliency, see http://technet.microsoft.com/en-us/library/dd638121.aspx.

Design assumptions and test environment user profile

Table 2 summarizes design assumptions, and the Exchange 2010 user profile used in both test environments.

Table 2. Simulated Exchange 2010 customer environment and requirements

Parameters	Details
Number of users per site	10,000 for each test environment
Exchange 2010 user profile	100 messages sent/received/day (0.10 IOPS)
Mailbox size	2 GB
Database read/write ratio	3:2 in Mailbox Resiliency configuration
1st test environment	• DAG for local and cross-site replication • 6 mailbox servers per DAG (4 local and 2 remote) • 3 database copies (2 local and 1 remote)
2nd test environment	• DAG for local and RecoverPoint for cross-site replication • 4 mailbox servers per DAG (local site) • 2 stand-alone mailbox servers at the DR site • 3 database copies (2 local with DAG, and 1 remote with RecoverPoint)
Simulated RPO	Local and remote – 5 min.
Simulated RTO	Local site – 15 min. DR site - 1 hr.
Cross-site network latencies	50, 150, 250 ms. tested
Connectivity between sites	1 Gb/s Ethernet and 155 Mb/ps (OC-3) tested
Storage connectivity to physical host	2 x 8 Gb/s FC ports
Disk type for Exchange databases and logs	2 TB SATA (7.2k rpm)
Disk type for RecoverPoint journals	600 GB FC (10k rpm)

Hardware and software components

Table 3 and Table 4 list the hardware and software components used to build our test environments.

Table 3. Hardware components

Item	Description
Storage platform	Two CLARiiON CX4-960 storage arrays (1 per site)
Storage connectivity to host (FC, iSCSI)	FC
Number of storage controllers (SPs)	Two per each CX4-960
Storage processor (SP) memory	32 GB (16 GB per SP)
Number of storage ports used	4 (2 per storage controller)
Fabric switch	8 Gb/s FC switching module
Physical hypervisor hosts	3 servers with quad eight-core Intel Xeon x 7560 processors @ 2.26 GHz CPU, 192 GB RAM
Host bus adapter (HBA)	8 GB HBAs (2 per hypervisor host)
Number of HBAs/host	2 per host
Total number of drives for Exchange database and logs tested in the solution	192 / 2 TB SATA, 7.2k rpm (96 drives per each test environment to support three database copies)
Total number of drives used for RecoverPoint journals (Production and DR site)	36 / 600 GB, 10k FC drives
Replication appliances	Four RecoverPoint Appliances GEN4, two per site
Distance emulator	Anue Gigabit Ethernet Multi-profile Network Emulator (GEM)

Table 4. Software components

Item	Description
CLARiiON operating environment	EMC FLARE® 30
RecoverPoint	Version 3.3 SP1
Multipath and I/O balancing software	EMC PowerPath®/VE 5.4.2 for VMware vSphere
Hypervisor operating system	VMware vSphere 4.1
Virtual machines operating system	Windows Server 2008 R2
E-mail software suite	Exchange Server 2010 SP1
Exchange storage performance validation tool	Exchange Jetstress 2010 64-bit, version 14.01.0225.017
Exchange server performance validation tool	Exchange Load Generator 64-bit, version 14.01.0180.003

Common design methodologies

The following sections describe the common design methodologies used during the build phase of the two simulated Exchange 2010 test environments. These include:

- Exchange 2010 mailbox resiliency with native DAG replication
- Exchange 2010 mailbox resiliency with native DAG at local site and cross-site resiliency by RecoverPoint

Virtualized Exchange 2010 using vSphere 4.1

In both test environments, we have virtualized Exchange 2010 using vSphere 4.1, and built a SAN architecture with the CLARiiON CX4-960 storage array. It is important to note that the same best practices and Exchange 2010 design guidelines still apply if these solutions utilized Microsoft Hyper-V as the virtualization platform.

Virtualization provides greater user resource consolidation with more flexible Exchange DR choices. Today, Microsoft fully supports Exchange 2010 on both the Hyper-V platform and VMware technology, as well as on all virtualization products that comply with the Microsoft server virtualization validation program (SVVP). At the time of this writing, for Exchange Server 2010, Microsoft supports all server roles in virtualized environments, except for the Unified Messaging role.

For more details about the SVVP program, see the Microsoft website, see http://www.windowsservercatalog.com/svvp.aspx.

For Exchange 2010 requirements in virtualization deployments, see: http://technet.microsoft.com/en-us/library/aa996719.aspx.

Building-block design for Exchange Mailbox virtual machines

Both test environments use the EMC proven building-block design methodology. A building-block represents a required number of server and storage resources per mailbox server that are scalable, and will guarantee best performance and optimal configuration based on the known Exchange user profile, and expected load. This methodology allows customers to simplify deployment of Exchange in their data centers and makes it easier to scale when additional users are added.

For more details on how to use the EMC building-block methodology for Exchange 2010 deployments, see chapter 4 in the *Zero Data Loss Disaster Recovery for Microsoft Exchange 2010—Enabled by EMC Unified Storage, EMC Replication Enabler for Exchange 2010, Brocade End-to-End Network, Dell Servers, and Microsoft Hyper-V* white paper available in the Resource library on www.EMC.com.

In both test environments we have designed Exchange mailbox virtual machines by using a building-block based on 5,000 users. The storage, CPU and memory resources for this building-block are derived from the design assumptions listed in Table 2. For additional Microsoft guidelines applied to our testing, see http://technet.microsoft.com/en-us/library/ee712771.aspx.

Exchange mailbox virtual machine building block details

Table 5 summarizes the Exchange mailbox virtual machine building-block for 5,000 users with a 100-message profile, and a 2 GB mailbox quota.

Table 5. Mailbox server building-block summary

Users per Exchange Server virtual machine	Disks per Exchange Server virtual machine	vCPUs per Exchange Server virtual machine	Memory per Exchange Server virtual machine
5,000	16 2 TB SATA disks in RAID 1/0 (for DBs and logs)	6	24 GB

Storage building-block details

Table 6 summarizes additional storage details for a mailbox server building-block with 5,000 users.

Table 6. Mailbox server storage building-block summary

Item	Value
Number of users	5,000 (in a switchover condition – 2 DAG copies)
User profile	0.10 IOPS
Mailbox size	2 GB
Disk type	2 TB SATA, 7.2k rpm
RAID type	RAID 1/0 storage pools, 16 spindles per pool
Raw storage capacity	29,349 GB
Formatted storage capacity to host	14,280 GB
Databases per server/users per database	8 databases, 625 users per database
LUNs/LUN sizes	16 LUNs: 8 DB LUNs - 1.7 TB 8 log LUNs - 85 GB

Exchange 2010 test configurations

Testing targets the following two configurations:

- 1st test environment—Exchange 2010 with cross-site DAG for local and remote replication
- 2nd test environment—Exchange 2010 with local DAG and RecoverPoint CRR for cross-site replication

Design details for Exchange 2010 configuration with cross-site DAG for local and remote replication

Figure 1 depicts the physical architecture of the test environment where Exchange 2010 is deployed in a cross-site DAG configuration. This environment is configured with Exchange 2010 DAG stretched across two physical sites. In this configuration, replication is performed over the IP network using built-in native Exchange 2010 continuous replication technology (log shipping). The DAG configuration in this solution is comprised of three copies—two database copies reside at the primary site, and a third copy resides at the DR site. Site activation and DR procedures in this configuration involve activating the third copy at the remote site.

For more information about data center switchovers and the reactivation process, see http://technet.microsoft.com/en-us/library/dd351049.aspx.

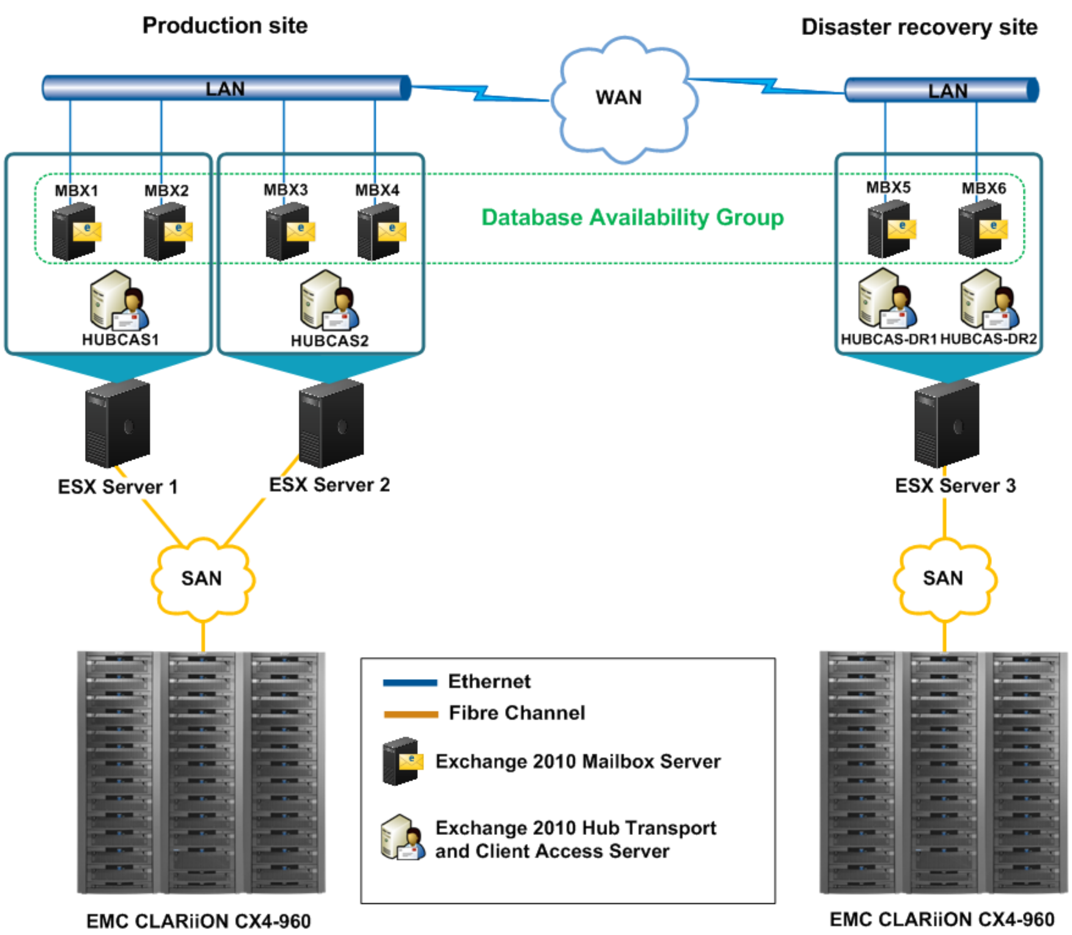

Figure 1. Exchange 2010 mailbox resiliency with native DAG replication (test environment 1)

Server roles in test environment 1

As illustrated in Figure 1, two VMware ESX® servers at the primary site are configured to host four Exchange 2010 Mailbox Servers, and two Hub Transport and Client Access role servers. One ESX server at the DR site is configured to host two mailbox servers and two hub transport servers. All six mailbox servers are configured to a single DAG spanning two physical sites.

Mailbox servers at the primary site were designed to host 5,000 users during local switchovers. A total of eight databases were configured per mailbox server with four primary databases copies (C1) hosting 2,500 users and four secondary copies (C2) hosting another 2,500 users. Mailbox server virtual machines at the DR site were designed to host the third database copy (C3) and to accommodate 5,000 users should a site failover occur.

DAG design for local and cross-site mailbox resiliency

Figure 2 provides details on how DAG copies are configured per each mailbox server. In this test environment, 10,000 users are configured across 16 databases with 625 users per database. All 16 databases are then balanced across six DAG member servers. Eight databases are configured for each mailbox VM, where four copies are active and four are passive.

This design provides redundancy and service availability during maintenance of the Mailbox server virtual machines, or an ESX server. For example, during maintenance on Exchange Mailbox Server MBX1, the secondary local copies (C2) are activated on another Mailbox Server MBX3 on a different ESX server. Furthermore, during maintenance of the physical hypervisor host (ESX1) both MBX1 and MBX2 mailbox servers will be unavailable. All C2 copies are automatically activated on mailbox servers MBX3 and MBX4, and configured on a different hypervisor host (ESX2). During a site failover event, mailbox servers MBX5 and MBX6, hosting third database copy (C3), will be activated using site activation process. This manual process can be scripted to reduce RTO.

	Production Site				DR Site	
	ESX 1		ESX 2		ESX 3	
	Database Availability Group					
	MBX1	MBX2	MBX3	MBX4	MBX5	MBX6
Database 1-4	C1		C2		C3	
Database 5-8	C2		C1		C3	
Database 9-12		C1		C2		C3
Database 13-16		C2		C1		C3

Figure 2. DAG design details for test environment 1

Design details for Exchange 2010 with local DAG and RecoverPoint CRR for remote replication

Figure 3 depicts the physical architecture of the test environment where Exchange is deployed in a local site DAG configuration, and RecoverPoint CRR is integrated to replicate database copies to a remote DR site. This test environment is configured with Exchange 2010 DAG at the local production site for HA (two copies). A third copy is created by replicating one of the local copies to a remote DR site using RecoverPoint CRR.

In this configuration, the DR procedure involves activating remote copies by using the Exchange 2010 database portability process. With this process users are re-homed to new databases on stand-alone mailbox servers that are predeployed at the DR site. The process involves simple procedures where all the data storage at the DR site is available to mailbox servers using RecoverPoint CRR. In this case, users are re-homed using Exchange PowerShell cmdlets. Replication Manager can also be integrated in to similar solutions.

Note. Although EMC Replication Manager was not in the scope of this configuration, we suggest you consider implementing it with all RecoverPoint solutions as it can help to manage automation of creating and mounting VSS snapshots (bookmarks) created by RecoverPoint. In addition, it can provide an option to automatically replay Exchange logs after the database is mounted.

The database portability procedures used during the failover process are described in detail later in this document, Exchange 2010 site failover process with RecoverPoint.

For more information on using the database portability process, see http://technet.microsoft.com/en-us/library/dd876873.aspx.

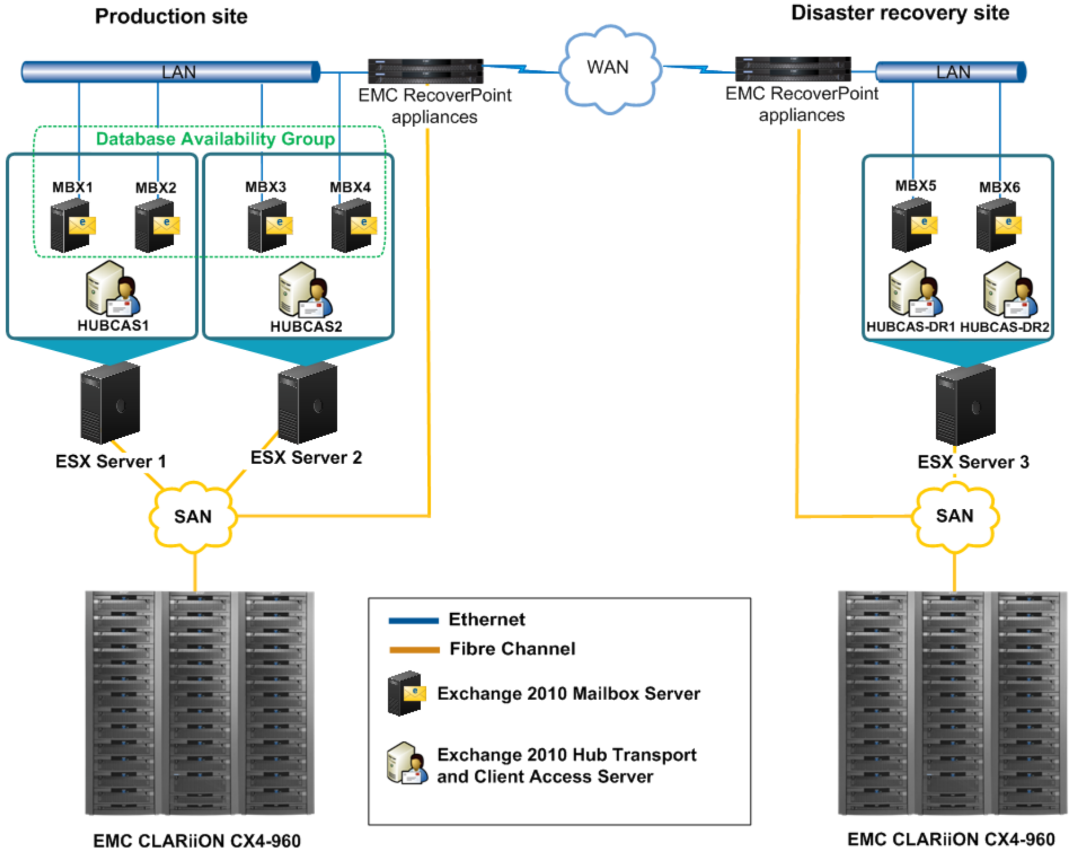

Figure 3. Native DAG at local site and cross-site resiliency with RecoverPoint CRR (test environment 2)

DAG design for local mailbox resiliency

Figure 4 provides greater detail on how database copies are configured for each mailbox server. The design is very similar to the first tested configuration, except that there are no DAG members at the remote site. Instead, there are two stand-alone mailbox servers that are configured to take over the load during the site failover by using the database portability process. During a site failover event, mailbox servers MBX5 and MBX6 will host the third database copy that is replicated by using RecoverPoint. In this case, the database portability process is used to re-home Exchange 2010 users to new mailbox servers. This manual process can be scripted to reduce RTO.

	Production Site				DR Site	
	ESX 1		ESX 2		ESX 3	
	Database Availability Group				Stand-alone Mailbox Servers	
	MBX1	MBX2	MBX3	MBX4	MBX5	MBX6
Database 1-4	C1		C2		C3	
Database 5-8	C2		C1		C3	
Database 9-12		C1		C2		C3
Database 13-16		C2		C1		C3

Figure 4. DAG design details for test environment 2

RecoverPoint CRR design

Our second tested DR option (local Exchange DAG and RecoverPoint for cross-site replication) is configured so that RecoverPoint replicates the active database copies from the primary production site to a secondary DR site. We have created a RecoverPoint CG for each database as this configuration provides better recovery granularity. Each database and its corresponding log files are configured as a single replication set and placed in to a single CG. The advantages of this design include:

- More granular backup and recovery options
  - Backup or restore can be performed on a single Exchange database
- Smaller bookmark sizes, which create smaller data lag
  - The smaller the data lag, the smaller the RPO
- Flexibility to prioritize initialization and replication per Exchange database level
  - If databases include different tiers of users, the SLA may require faster RTO for higher tier users
  - Policy settings can be adjusted per CG to accommodate a specific protection window (as needed)
- Greater flexibility to load balance between busy RecoverPoint Appliances (RPAs). CGs can be moved to an underutilized RPA

Note Review the *EMC RecoverPoint Administrator's Guide* at http://powerlink.emc.com for best practices on configuring RecoverPoint.

Configuring RecoverPoint CGs

Table 7 and Figure 5 illustrate how RecoverPoint CGs were configured in the test environment for a single Exchange mailbox server with eight databases. In our configuration we have selected only the active primary copies for replication. This will provide no data lag in the DR copies. It is very important that the Exchange administrator selects the active copy when adding a new Exchange database to a CG.

Note If your requirements are to replicate both local Exchange 2010 copies, you need to consider providing an additional storage at the DR site for that copy.

Table 7. Example of RecoverPoint CG configuration for one mailbox server

Replication set name	Replication set members (LUNs)
DB1	MBX1_DB1 & MBX1_LOG1
DB2	MBX1_DB2 & MBX1_LOG2
DB3	MBX1_DB3 & MBX1_LOG3
DB4	MBX1_DB4 & MBX1_LOG4
DB5	MBX3_DB5 & MBX3_LOG5
DB6	MBX3_DB6 & MBX3_LOG6
DB7	MBX3_DB7 & MBX3_LOG7
DB8	MBX3_DB8 & MBX3_LOG8

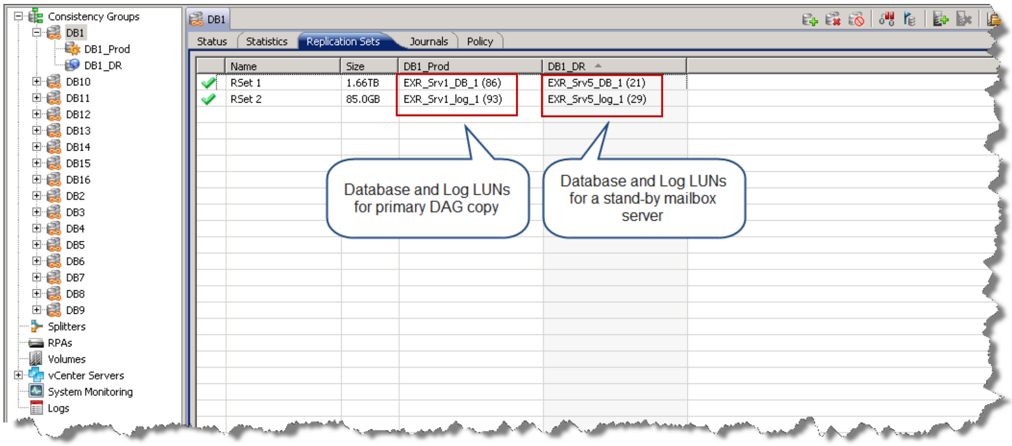

Figure 5. RecoverPoint CG configuration

Configuring the RPO policy with RecoverPoint

There are distinct advantages to configuring the RPO at the CG level with RecoverPoint. This protection policy can be optimized according to specific requirements, or can be regulated by RecoverPoint. For remote replication, RecoverPoint can:

- Utilize as much bandwidth as necessary to keep the RPO to a minimum
- Use more of the available bandwidth (when required) to keep the RPO within its allowed range.

For example, in our configuration, where each database is in its own CG, we can set different policies to provide a different protection level for each Exchange database. This can be used when different tiers of users (regular users, executive users, etc.) are configured in the environment.

Figure 6 provides details for multiple RPO options that can be set at the individual CG level. The RPO can be set at multiple levels from GB down to an individual write that is being sent to a storage array.

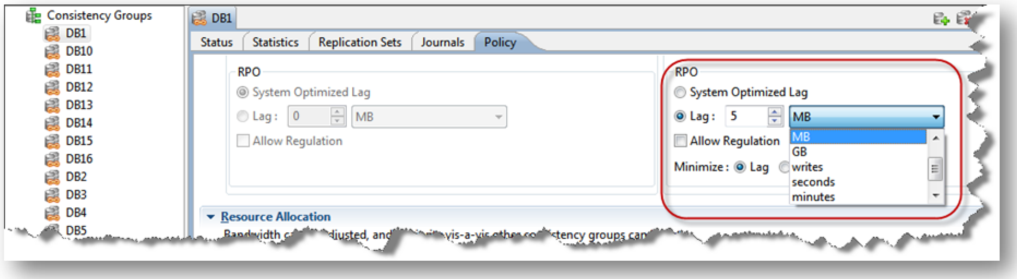

Figure 6. RPO choices per RecoverPoint CG

Automatic periodic bookmarking (snapshots)–group sets

A group set is a set of CGs to which the system applies parallel bookmarks at a user-defined frequency. A group set allows you to automatically bookmark a set of CGs so that the bookmark represents the same recovery point in each CG in the group set. This allows you to define consistent recovery points for CGs distributed across different RPAs. Group sets are useful for CGs that are dependent on one another or that must work together as a single unit.

In our configuration we have configured four group sets, where each group set consists of four CGs, each representing active database copies per each Exchange server. This configuration provides for easy recovery as all bookmarks for every CG within a group set will have the same RPO and time stamp.

Figure 7 provides details on configuring group sets for four mailbox servers.

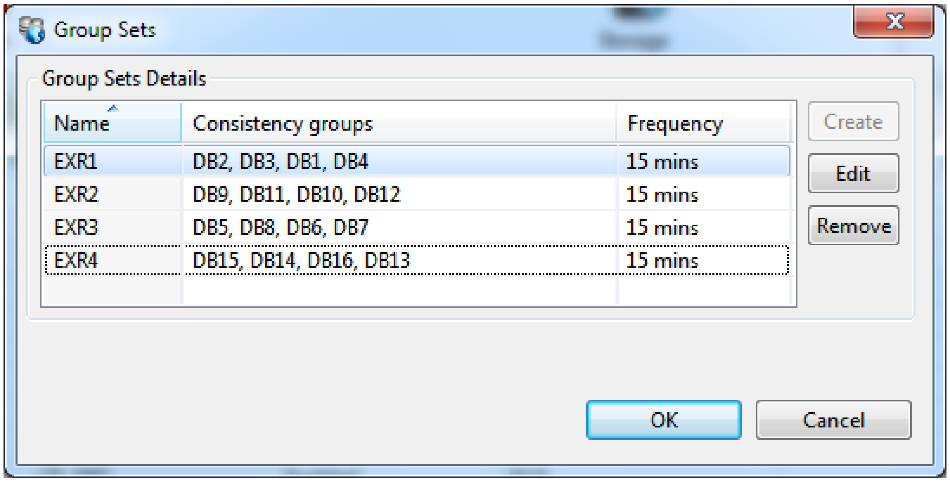

Figure 7. RecoverPoint group sets

RecoverPoint journals

RecoverPoint journal volumes store consistent point-in-time snapshots (or bookmarks) to enable point-in-time recovery of Exchange data. These bookmarks allow DVR-like rollbacks of the data to different points in time. There are two journal volumes required per each CG in a RecoverPoint CRR configuration—one at each replication site.

Because we are replicating 16 Exchange databases in our test configuration, we have configured 16 RecoverPoint CGs, one group per Exchange database and its transaction logs. This totals 32 journal volumes (16 at each site).

Considerations when sizing journal volumes

The two most important considerations when sizing storage for journal volumes are:

- Performance
- Protection window

For optimum journal performance, make sure to choose the appropriate RAID and drive types. In our testing we used 600 GB 10k rpm drives in a RAID 1/0 configuration for journal volumes.

The protection window depends on the change rate of the application and how far back Exchange administrators want to go to capture snapshots (bookmarks). Both of these factors will determine the required journal size.

For full details, including the formula for calculating journal volume size, see the *EMC RecoverPoint Administrator's Guide* available on www.powerlink.emc.com.

RecoverPoint storage requirements

Table 8 provides a summary of the RecoverPoint storage used in testing. In addition to replication volumes at each site, RecoverPoint requires a very small volume, called the repository volume that must be dedicated on the SAN-attached storage at each site, for each RPA cluster. It stores configuration information about the RPAs and CGs, which enables a properly functioning RPA to seamlessly assume the replication activities of a failing RPA from the same cluster.

For our RecoverPoint CRR test configuration we have configured small 20 GB journal volumes per each CG at the primary site, as they are only used during failback from DR site. We also configured 200 GB journal volumes for each CG at DR site based on a 20 percent change rate as a rule of thumb. Note that the journal volume sizes we chose (see Table 8) are configured to accommodate the multiple test scenarios and may differ from guidelines listed in the *EMC RecoverPoint Administrator's Guide*.

Note EMC strongly advises that you follow the RecoverPoint guidelines outlined in the *EMC RecoverPoint Administrator's Guide* to ensure that the journal volumes meet your specific RPO requirements.

Table 8. RecoverPoint volumes breakdown

Volume type	Production site	DR site	Comments
Repository	3 GB	3 GB	One volume per storage array
Journals	20 GB	200 GB	Per CG at each site (16 per site)
Replication	N/A	1.7 TB for DB volume 85 GB per log volume	Replication volume sizes must be equal production databases and log volumes sizes

Test methodology

When designing any DR solution it is important to understand the capacity, speed, and utilization of the link between the production and DR sites. In many cases, the link is the factor that causes a solution to fail. To better understand how network latencies and bandwidth can impact Exchange DR replicated solution, we have performed multiple replication tests in both environments and observed the timing and link utilization. The results can be used as a reference point for sizing a DR solution that meets your organization's specific RTO and RPO requirements.

Distance emulator

In both tested Exchange environments, the Anue Gigabit Ethernet Multi-profile Network Emulator (GEM) is configured to inject latencies and to simulate different network type by reducing bandwidth. This device was added on the replication network between production and DR sites to validate database seeding performance with cross-site DAG and to validate RecoverPoint replication performance at varying distances. More information about Anue network emulators can be found at http://www.anuesystems.com.

Understanding network latencies

Generally speaking, network latencies correspond to different distances. These round-trip latencies can vary from 0 ms to 250 ms and higher. In our testing, we used the following two link types:

- OC-3 with 155 Mbps throughput
- 1 Gb/s Ethernet

Table 9 details the correlation between network distance, and its corresponding average latencies.

Note Numbers represented in Table 9 show estimates and will vary depending on a number of factors including the quality of the link, network hops, and backbone hops.

Table 9. Average link latencies

Approximate distance	One-way latency	Round-trip latency
50 km	0.25 ms	0.5 ms
500 km	2.5 ms	5 ms
1,000 km	5 ms	10 ms
5,000 km	25 ms	50 ms
1,000 km	50 ms	100 ms
20,000 km	100 ms	200 ms
22,500 km	125 ms	250 ms

Exchange data used for replication

The Microsoft Exchange Load Generator tool (LoadGen) was used to produce the data used in both test environments. Because of the issue with excessive content index (CI) size produced by this tool, we had to initialize mailboxes smaller than 2 GB as was originally designed for our test configuration to accommodate for an extra space. Table 10 details the data used in replication.

Table 10. Replicated Exchange data summary

What's replicated?	Database and content index data	Log data	Total replicated data
Single Exchange database	800 GB (500 GB DB size + 300 GB CI size)	5 GB	805 GB
Four Exchange databases (one mailbox server)	3,200 GB (4 x 800 GB)	20 GB	3,220 GB

Note An excessive CI size seen in the test environment (over 50 percent of the database size) is due to the Dynamic Content Generation feature introduced in the LoadGen 2010 tool. In production Exchange 2010 environments, CI size is usually 10 percent of the database size.

For more information about LoadGen deployment and usage, see http://technet.microsoft.com/en-us/library/dd335108.aspx.

Database seeding process

Before Exchange 2010 mailbox resiliency in DAG environment is fully deployed, all database copies must be replicated to each DAG member. This process is known as database seeding, in which a copy of a mailbox database is added to another mailbox server. This process is significant because each replicated database becomes the baseline database for the copy, and is used during a switchover or failover condition. The copy must be readily available, and show a "healthy" status, to be used for activation.

Database seeding is the most time consuming, yet critical process that requires careful planning when deploying Exchange 2010 in a DAG configuration. Seeding is required under the following conditions outlined in the article provided by Microsoft at http://technet.microsoft.com/en-us/library/dd351100.aspx. Below are some of the highlights from the article:

- When a new passive copy of a database is created, seeding can be postponed for a new mailbox database copy. But eventually, each passive database copy must be seeded in order to function as a redundant database copy.
- Sometimes a failover occurs in which data is lost as a result of the passive database copy having become diverged and unrecoverable.
- When the system has detected a corrupted log file that cannot be replayed into the passive copy of the database.
- After an offline defragmentation of any copy of the database occurs.
- After the log generation sequence for the database has been reset back to 1.

Note When using RecoverPoint for replication there is no seeding or slivering operation that impacts the production databases. The process is performed in the background.

Possible causes for seeding to suddenly terminate

Updating a database copy can take a very long time, especially if the database is very large, or due to high network latency and low network bandwidth. A database copy can be seeded using either the active copy or an up-to-date passive copy as the source for the seed. When seeding from a passive copy, be aware that the seed operation will terminate with a network communication error under the following circumstances:

- If the status of the seeding source copy changes to "Failed" or "Failed And Suspended"
- If the database fails over to another copy

Recommendations on seeding database copies simultaneously

It is possible to seed multiple database copies simultaneously, as we did during testing of this solution. However, Microsoft recommends that you seed only the database file, and omit the content index catalog. Use the *`DatabaseOnly`* parameter with the Update-MailboxDatabaseCopy cmdlet. If you do not use the *`DatabaseOnly`* parameter when seeding multiple targets from the same source, the task will fail with `SeedInProgressException` error FE1C6491. For more details, see http://technet.microsoft.com/en-us/library/dd351100.aspx.

Note We did not see the above `SeedInProgressException` error during our testing, and were able to seed the database and content index simultaneously.

Validation and test results

Test results presented in this paper include:

- Data replication performance
  - Initial database seeding performance in a cross-site DAG at multiple latencies and link types
  - Initial RecoverPoint synchronization ("full sweep") between two sites at multiple latencies and links types
  - Replication process under load-capturing network utilization
- DR failover process
  - Functional failover process that uses the data center activation process with cross-site Exchange 2010 DAG, while capturing RPO and RTO
  - Functional failover process that uses the Exchange 2010 database portability feature with RecoverPoint CRR, while capturing RPO and RTO
- Storage validation for both test environments (see Appendix B)
  - Storage building-blocks from two environments are combined to validate performance of the CX4-960 with 2 TB SATA drives (20,000 users with over 55 TB of storage)

Data replication performance results

Initial replication

Table 11 summarizes replication performance results observed during the database seeding process with Exchange DAG, and the first time initialization process with RecoverPoint, also known as a "full sweep." For more details about different replication test iterations, see Appendix A.

Results observed from replication testing show that the RecoverPoint replication process is more than 10 times faster than the native host-based DAG replication method during replication of 3.2 TB of Exchange data (four databases replicated simultaneously) over a 1 Gb/s network with 150 ms latencies. The RecoverPoint initial replication process (full-sweep) was completed within six hours, where host-based DAG replication completed over a 12-hour window. This significant performance gain was possible due to 5x data compression and efficient utilization of the network link provided by RecoverPoint. It is also important to note that "sweep" time operates in the background which enables the use of the Exchange server while the DAG seeding forces you to wait. Although it took six hours for RecoverPoint to replicate the data, there is no impact on production or backup operations.

Note The following performance results are based on the test lab network configuration, and can vary in your production environment depending on factors like the quality of the link, network hops, and backbone hops.

Table 11. Initial data replication performance results

Link type and network latency	DAG replication performance results		RecoverPoint replication performance results	
	Elapsed seeding time	Network utilization	Elapsed full sweep time	Network utilization
1 Gb/s , 0 ms	12 hrs, 30 min	50 %	6 hrs	50 %
1 Gb/s , 50 ms	18 hrs	12 %	8 hrs	45 %
1 Gb/s , 150 ms	120 hrs (5 days)	10 %	13 hrs	40 %
1 Gb/s , 250 ms	192 hrs (8 days)	5 %	22 hrs	28 %
OC-3 , 150 ms	126 hrs (5.25 days)	15 %	15 hrs	40 %

Another significant RecoverPoint advantage is observed when 250 ms latency was injected on the replication network between the primary and DR sites. A single Exchange server with four databases and 3.2 TB of storage took only 22 hours to replicate, compared to the native Exchange replication method which took eight days to replicate.

Important! Based on observed results, customers deploying Exchange 2010 in a DAG should pay attention to sizing their log LUNs. By doing so they can accommodate for all logs that will be accumulated during prolong database seeding time periods. This is necessary to avoid database dismount due to log LUNs overflow as backups cannot be performed during seeding time. This can significantly affect RPO requirements. At the same time, this should not be a concern if RecoverPoint CRR is used for replication, because backups can be performed without any restrictions during replication process.

Replication under load

To determine how Exchange replication impacts the network between production and DR sites, we performed a test that simulated normal Exchange 2010 operating condition. We used LoadGen to simulate Exchange 2010 workload for 10 hours while monitoring network utilization. We performed this test under 150 ms round-trip latency injected on the 1 Gb/s link between sites.

Table 12 shows results observed during the test. We saw that DAG replication used 30 percent of the link, where RecoverPoint used only 15 percent of the bandwidth. This performance gain was possible due to 5 to 6 times RecoverPoint data compression and efficient utilization of the network link.

Table 12. Network utilization during replication

Link type and network latency	Network utilization with DAG replication	Network utilization with RecoverPoint replication
1 Gb/s , 150 ms	30%	15%

Exchange 2010 site failover process with cross-site DAG

A data center or site failure is managed differently from the types of failures that can cause a server or database failover. For local server failures with a HA configuration, automatic recovery is initiated by the system. This failure typically leaves the messaging system in a fully functional state.

In contrast, a data center failure is considered a DR event. Recovery must be manually performed and completed for client service restoration, and for the outage to end. The process is referred to as a data center switchover. As with many DR scenarios, prior planning and preparation for a data center switchover can simplify the recovery process, and reduce the duration of the outage.

The data center switchover process is covered in great detail in the following article, http://technet.microsoft.com/en-us/library/dd351049.aspx.

Reviewing the data center switchover process

Here are the four basic, high-level steps used to perform a data center switchover after making the initial decision to activate the second data center:

- Terminate the partially running data center
- Validate and confirm the prerequisites for the second data center
- Activate the mailbox servers
- Activate the other server roles (Hub, CAS)

How to execute a data center switchover

Use the following steps to perform a data center switchover.

1. Terminate any surviving DAG members in the primary data center. Type the appropriate command, as follows:

   a. If servers in the primary site are still accessible:

   ```
   Stop-DatabaseAvailabilityGroup -ActiveDirectorySite
   <insertsitename>
   ```

   b. If servers in the primary site are not accessible:

   ```
   Stop-DatabaseAvailabilityGroup -ActiveDirectorySite
   <insertsitename> -ConfigurationOnly
   ```

   Important! Failure either to turn off the mailbox servers in the failed data center or to successfully perform the `Stop-DatabaseAvailabilityGroup` command against the servers will create the potential for "split-brain syndrome" to occur across the two data centers.

2. Validate and confirm the prerequisites for the secondary data center. The purpose of this step is to inform the servers in the second data center about which mailbox servers are available to use when restoring service. The second data center must now be updated to represent the primary data center servers that have stopped.

   a. Run the following cmdlet in the second data center:

   ```
   Stop-DatabaseAvailabilityGroup -ActiveDirectorySite
   <insertsitename> -ConfigurationOnly
   ```

3. Activate the mailbox servers at the secondary data center.

   a. Stop the Cluster service on each DAG member in the second data center. You can use the `Stop-Service` cmdlet to stop the service (for example, `Stop-Service ClusSvc`), or use `net stop clussvc` from an elevated command prompt.

   b. Activate the mailbox servers in the standby data center by running the following power shell cmdlet:

   ```
   Restore-DatabaseAvailabilityGroup -
   ActiveDirectorySite <insertsitename>
   ```

   Note When this command completes successfully, the quorum criteria reduces to the servers in the standby data center. If the number of servers in that data center is an even number, the DAG switches to the alternate witness server, as identified by the setting on the DAG object.

   c. Activate the databases by running the following power shell cmdlet:

   ```
   Get-MailboxDatabase <insertcriteriatoselectDBs> |
   Move-Active-mailboxDatabase -ActivateOnServer
   <DAGMemberinSecondSite>
   ```

   Or, use the following command:

   ```
   Move-Active-mailboxDatabase -Server
   <DAGMemberinPrimarySite> -ActivateOnServer
   <DAGMemberinSecondSite>
   ```

   d. Check the event logs and review all error and warning messages to ensure that the secondary site is healthy. Follow up and correct all issues prior to mounting the databases.

   e. Mount the databases using the following power shell cmdlet:

   ```
   Get-MailboxDatabase <DAGMemberInSecondSite> | Mount-
   Database
   ```

4. Activate the other server roles at the secondary data center.

   Activating the Internet-facing Client Access servers involves changing DNS records to point to the new IP addresses that are configured for the new service endpoints. Clients will then automatically connect to the new service endpoints in one of two ways:

   - Clients can continue to try to connect, and should automatically connect after the TTL has expired for the original DNS entry, and after the entry is expired from the client's DNS cache. Users can also run the `ipconfig/flushdns` command from a command prompt to manually clear their DNS cache.
   - Outlook clients starting or restarting will perform a DNS lookup on startup and will get the new IP address for the service endpoint, which will be a Client Access Server or array in the second data center.

Validating the data center switchover with LoadGen

The following steps detail the data center switchover using LoadGen.

- Change the DNS entry for the Client Access Array to point to the VIP of the HWLB in the secondary site.
- Run the `ipconfig /flushdns` command on all LoadGen servers.
- Restart the LoadGen load.
- Verify that the CAS servers in the secondary site are now servicing the load.

Data center failback procedure

Use these steps to failback the mailbox server role.

Note The mailbox server role should be the first role to failback to the primary data center.

1. Run the following power shell cmdlet to re-incorporate the DAG members into the primary site.

```
Start-DatabaseAvailabilityGroup ActiveDirectorySite
<insertsitename>
```

2. Verify the state of the database copies in the primary data center.

   Important! After the mailbox servers in the primary data center are incorporated into the DAG, allow them to synchronize their database copies. Depending on the nature of the failure, the outage length, this may require reseeding the database copies.

3. Reconfigure the DAG to use a witness server in the primary data center using the following power shell cmdlet. During the data center switchover process, the DAG was configured to use an alternate witness server.

```
Set-DatabaseAvailabilityGroup -Identity <DAGName> -
WitnessServer <WitnessServer>
```

4. Dismount the databases that you are reactivating in the primary data center from the secondary data center. Use the following power shell cmdlet:

```
Get-MailboxDatabase <insertcriteria> | Dismount-Database
```

5. After dismounting the databases, move the Client Access Server URLs from the secondary data center to the primary data center by changing the DNS record for the URLs to point to the Client Access server or to the array in the primary data center.

The data center switchover steps described above for the Exchange 2010 DR process can be used as an example for measuring RTO during Exchange DR planning phase. The duration of each step can be different in real production environments. In our test lab environment it took just under one hour to failover four mailbox servers (10,000 users) to a DR site using data center switchover process. The time it took to failover was within our simulated RTO requirements set for this configuration.

Exchange 2010 site failover process with RecoverPoint

The site failover process with RecoverPoint requires access to the RecoverPoint management interface located at either the production or DR sites. The process is described in more detail in the *EMC RecoverPoint—Replicating Microsoft Exchange Server* technical notes available on http://powerlink.emc.com.

The detailed best practices outlined in the above document enable Exchange and system administrators to use RecoverPoint failover and failback capabilities to support database recovery for both standalone and HA deployments of Exchange 2010 mailbox servers.

1. For a planned site failover, you must first shut down the Exchange servers at the primary site.
2. Take a parallel bookmark of the Exchange servers (or, all CGs for the appropriate server).
3. Enable image access at the DR site.
4. Power-on the DR Exchange Server VM (if not already turned on) and verify that the disks are online.
5. Use a stand-by Exchange server VM at DR site to verify that the DBs are in the clean shutdown state.

Note The databases will be in the “dirty shutdown” state unless they were shutdown gracefully before the RecoverPoint bookmark is taken. With a graceful shutdown and bookmark taken after, the status will be a “clean shutdown” state.

Example

```
eseutil /mh c:\MP\DB1\DBR01.edb
```

6. Run `softrecovery` to replay the uncommitted logs into the database. This step can be skipped if verification from step 5 confirmed a “clean shutdown” state of the databases.

   Example

   ```
   C:\MP\LOG1>eseutil /r E01
   ```

7. Create new mailbox database instances on the destination server at DR site.

   Example

   ```
   New-MailboxDatabase -Name DB1_DR -Server EXR-SRV5 -
   EdbFilePath "c:\mp\db1\RP\DBR01.edb" -LogFolderPath
   "c:\mp\log1\RP
   ```

8. Enable the restore flag on the newly created mailbox databases.

   Example

   ```
   set-MailboxDatabase DB1_DR -AllowFileRestore:$true
   ```

9. Move the replicated databases to a newly created database path, and mount them.

   Example

   Example: `Mount-Database DB1_DR`

10. Re-home users (AD configuration change only).

   Example

   ```
   get-Mailbox -Database DBR01 |where {$_.ObjectClass -
   NotMatch'(SystemAttendantMailbox|ExOleDbSystemMailbox)'}
   | Set-Mailbox -Database DB1_DR
   ```

11. Update the DNS and verify user e-mail access.

Reviewing Exchange DR failover tasks and execution times with RecoverPoint

Table 13 provides additional details for each recovery step performed during the Exchange DR process. Please note that the actual duration of each step is different in real production environments. Use durations provided here as reference points only.

In our test lab environment it took just under one hour to failover four mailbox servers (10,000 users) to a DR site using RecoverPoint with the Exchange 2010 database portability process using manual steps. The same process using scripted tasks is completed in approximately 20 minutes.

Table 13. Exchange failover tasks with RecoverPoint CRR

Task number	Task description	Per server (in min)	Per site (in min)
1	Shut down Exchange servers at the primary site for a planned failover.	1	3
2	Take a parallel bookmark of each Exchange server (scripted or manual).	0.5 (scripted) 1 (manual)	0.5 (scripted) 2 (manual)
3	Present the replica to the destination Exchange host. In the RecoverPoint GUI, enable image access by selecting a bookmark (snapshot) for recovery. If possible, use a RecoverPoint crash-consistent bookmark first. If a suitable RecoverPoint bookmark is not available, use the latest VSS application-consistent bookmark created by KVSS or Replication Manager (scripted).	.5 (scripted) 3 (manual)	.5 (scripted) 10 (manual)
4	On the DR Exchange Server verify that the disks are online and mount replicated volumes. The drive letter does not need to be identical to the drive letter that is on the source server. **Note:** If the presented replica volume could not be brought online using the Disk Manager, reboot the host and retry bringing the replica volume online and then assign mounting point.	1	3
5	Determine the replicated database's consistency state on destination host. Run the Exchange eseutil /mh command to determine the replicated mailbox database's consistency state. Attempting to mount a database in any state other than clean shutdown will fail.	1	3
6	Run softrecovery to replay the uncommitted logs into the database if the database was not in the clean shutdown state before mounted RecoverPoint bookmark was taken. Run the Exchange eseutil /r command to replay the "uncommitted logs," which eventually sets the state of the Mailbox database to clean shutdown and allows database to be mounted. Duration of softrecovery can depend on multiple factors such as logs outstanding and storage performance.	N/A	N/A
7	Create a directory under the root for the database and one for the logs. Create a new mailbox database instance on the destination server.	0.5 (scripted) 1 (manual)	0.5 (scripted) 10 (manual)

Task number	Task description	Per server (in min)	Per site (in min)
8	Use PowerShell cmdlets to set the destination mailbox database on destination host, and to enable restore (replacement) of replicated database files.	0.5 (scripted) 1 (manual)	0.5 (scripted) 5 (manual)
9	Dismount the database and place the replicated database file on the new (and empty) mailbox database path. On destination host verify that you have placed the replicated EDB file to the empty folder of the new database.	0.5 (scripted) 2 (manual)	0.5 (scripted) 10 (manual)
10	Mount the newly replicated mailbox database for the first time, so the EDB file is created. On destination host use PowerShell cmdlets to mount new mailbox database.	0.5 (scripted) 1 (manual)	0.5 (scripted) 5 (manual)
11	Re-home users by swapping Active Directory user accounts links to the Exchange mailboxes from the source server to the new destination Exchange server at DR site.	5	5
	Total duration in minutes	11 min scripted 17 min manual	19 min scripted 56 min manual

Conclusion

In this white paper we have discussed two methods that can be used with Exchange 2010 to provide cross-site resiliency and HA:

- Exchange 2010 mailbox resiliency with native DAG replication
- Exchange 2010 mailbox resiliency with native DAG at local site and cross-site resiliency by RecoverPoint

We have reviewed the design and configuration details for both options. We have also performed multiple tests in both test environments to validate the best possible and optimal replication option that will provide the smallest RPO and RTO possible.

Based on these test results, we observed that RecoverPoint CRR can provide advantages to native Exchange host-based replication when deployed in cross-site configurations.

- With RecoverPoint replication, Exchange administrators do not need to worry about not meeting nightly backup windows as the cross-site replication process will not interfere with local site DAG replication.
- With RecoverPoint replication, customers can use the point-in-time rewind feature that helps recover lost or damaged data and databases more quickly and easily, complementing any periodic backup or snapshot solution.
- RecoverPoint block-level replication significantly speeds deployment for large Exchange 2010 databases, which decreases replication time.
- Using RecoverPoint for cross-site replication simplifies Exchange 2010 DR operation, and allows customers to reduce RTO while achieving RPO during the database seeding process.

References

Refer to the following documents for additional information.

White papers

For white papers that provide similar solutions, see the following:

- *EMC RecoverPoint Family Overview—A Detailed Review*
- *Improving Microsoft Exchange Server Recovery with EMC RecoverPoint—A Detailed Review*

Product documentation

For additional information on the products discussed in this white paper, see the following:

EMC:

- *EMC RecoverPoint Administrator's Guide*

Microsoft:

- *Exchange Server 2010* article at http://technet.microsoft.com/en-us/library/bb124558.aspx

Next steps

To learn more about this and other solutions, contact an EMC representative or visit www.EMC.com.

Appendix A – Multiple replication test scenarios

This appendix provides details and performance results for the three test scenarios used to replicate the Exchange data using the native DAG host-based replication method and RecoverPoint's block-level replication method.

The three test scenarios include:

- Test scenario 1—Single database seeding
- Test scenario 2—Seeding one mailbox server
- Test scenario 3—Single database replication with RecoverPoint CRR (full sweep)

Table 14 lists details about the replicated data.

Table 14. Exchange replicated data summary

What's replicated?	Database and content index data	Log data	Total replicated data
One database	800 GB (500 GB-DB, 300 GB-CI)	5 GB	805 GB
Four databases	3,200 GB (4 x 800 GB)	20 GB	3,220 GB

Test scenario 1—Single database seeding

In the first round of testing we validated the database replication performance across sites using the native Exchange 2010 solution. Test scenario 1 is divided into four individual sub-tests to identify the seeding performance of a single database over various network latencies, as follows:

- **Test 1a**—Single database replication (local)
- **Test 1b**—Single database replication (cross-site at 50 ms latency)
- **Test 1c**—Single database replication (cross-site at 150 ms latency)
- **Test 1d**—Single database replication (cross-site at 250 ms latency)

Test scenario 1 results

Table 15 details results observed during seeding of a single database over multiple latencies.

Table 15. Test results for single database seeding with multiple network latencies

Test	Network round trip latency	Elapsed time	Throughput rate GB/min	Bandwidth utilization (%) 1 Gb/s link
Test 1a	0 ms (local)	6 hours	2.24 GB/min	20 %
Test 1b	50 ms	8 hrs, 15 min	1.63 GB/min	10-12 %
Test 1c	150 ms	19 hrs, 30 min	0.69 GB/min	8-10 %
Test 1d	250 ms	32 hours	0.042 GB/min	5-6 %

Test analysis (test 1a through test 1d)

As shown in Table 15, the network latency greatly impacts the duration of the database seeding process. The replication time increases from 6 hours to 32 hours as the network latency increases from 0 to 250 ms.

It is critical that the Exchange 2010 DAG design under these conditions is fully understood as logs truncation during databases seeding process is not allowed. This means that the Exchange administrator needs to factor in high latency conditions when sizing their log volumes and backup solution. Also, if there is any interruption on the network link during the seeding process, the administrator must restart seeding from the beginning.

Test scenario 2—Seeding one mailbox server

The second test validates the performance seeding of four databases for a single mailbox server. Four databases representing 2,500 users with a 3,220 GB data to be replicated.

Test scenario 2 is divided into four individual sub-tests to identify the seeding performance of a single mailbox server over various network latencies, as follows:

- **Test 2a**— Replication of four databases (local)
- **Test 2b**— Replication of four databases (cross-site at 50 ms latency)
- **Test 2c**— Replication of four databases (cross-site at 150 ms latency)
- **Test 2d**— Replication of four databases (cross-site at 250 ms latency)

Test scenario 2 results

Table 16 details the performance seeding of four databases for a single mailbox server.

Table 16. Test results for a single mailbox server seeding with multiple network latencies

Test	Network round trip latency	Elapsed time	Throughput rate GB/min	Bandwidth utilization (%) 1 Gb/s link
Test 2a	0 ms (local)	12 hrs, 30 min	4.27 GB/min	45-50 %
Test 2b	50 ms	18 hrs	2.98 GB/min	10-12 %
Test 2c	150 ms	120 hrs (5 days)	0.44 GB/min	8-10 %
Test 2d	250 ms	192 hrs (8 days)	0.28 GB/min	3-5 %

Test analysis (test 2a through test 2d)

As shown in Table 16, the network latency greatly impacts the duration of the database seeding process, especially when multiple databases are seeded simultaneously. The replication time increases from 12.5 hours to 8 days as the network latency increases from 0 to 250 ms.

Test scenario 3—Initial replication with RecoverPoint (full sweep)

In the last round of testing we validated RecoverPoint replication performance across sites. Test scenario 3 is divided into three individual sub-tests to identify the replication performance of a single database over various network latencies, as follows:

- **Test 3a**— Single database initial replication (cross-site at 50 ms latency)
- **Test 3b**—Single database initial replication (cross-site at 150 ms latency)
- **Test 3c**—Single database initial replication (cross-site at 250 ms latency)

Test scenario 3 results

Table 17 details replication performance results with RecoverPoint.

Table 17. RecoverPoint test results for a single mailbox server initial replication with multiple network latencies

Test	Network round trip latency	Elapsed time	Throughput rate GB/min	Bandwidth utilization (%) 1 Gb/s link
Testb3a	50 ms (local)	8 hrs, 30 min	6.13 GB/min	Avg 40-45 %
Test 3b	150 ms	13 hrs, 30 min	3.97 GB/min	Avg 35-40 %
Test 3c	250 ms	22 hrs, 20 min	2.40 GB/min	Avg 22-28 %

Test analysis (test 3a through test 3c)

Table 17 provides a RecoverPoint performance summary observed during seeding of a single database over various network latencies. As the results show, the network latency has a moderate impact on the duration of the database replication process. The replication time increases from 8.5 hours to 22.2 hours as network latency increases from 50 to 250 ms.

With RecoverPoint replication, Exchange administrators do not need to worry about not meeting nightly backup windows as the cross-site replication process will not interfere with local site DAG replication.

Appendix B – Exchange storage validation with Jetstress

The following sections describe the process of validating the storage configuration and performance levels of the Exchange 2010 building-blocks in the two test environments deployed on the CX4-960 storage array.

Microsoft Jetstress 2010

We used Microsoft Jetstress 2010 to validate our storage design, and verify the performance and stability of the CX4-960 storage array with 2 TB SATA drives. Jetstress helps verify storage performance by simulating an Exchange I/O load. Specifically, Jetstress simulates the Exchange database and log file loads produced by a specified number of users.

Additional information about Jetstress is available at http://technet.microsoft.com/en-us/library/dd335108.aspx.

Jetstress test configuration

We have combined storage configurations from both test environments and ran Jetstress using four mailbox servers, each supporting a 5,000 user load with a 0.10 IOPS user profile. Table 18 details the configuration used during Exchange 2010 storage validation with Jetstress.

Table 18. Simulated Exchange 2010 configuration for Jetstress testing

Item	Value
Total users	20,000 (from both test environments)
User profile	0.10 IOPS (100 messages/user/day)
Mailbox size	2 GB
Mailbox servers simulated	4
Users per server	5,000
Databases per server	8
Users per database	625
Disk type	2 TB, 7.2k rpm SATA drive
RAID type	RAID 1/0 storage pools
Disks per server and for total configuration	16 (64 total)
LUNs per server	16 LUNs (8 DB and 8 log)
Number of DAGs and database copies simulated	1 DAG with 3 copies

Jetstress test results—total IOPS achieved on a per server basis

Figure 8 and Table 19 represent the total Jetstress IOPS results achieved on a per server basis. Information includes:

- The sum of all of the achieved transactional IOPS
- The average of all the logical disks' I/O latency in the two-hour Jetstress test

The two-hour Jetstress test produced 2,181 achieved transactional Exchange 2010 user IOPS. Note that achieved transactional IOPS value in the Jetstress report include only database reads and database writes IOs, but *do not* include Background Database Maintenance (BDM) I/Os, log I/Os and log replication I/Os.

The same two-hour test also produced 4,000 total Exchange IOPS (logged at the host to storage), which includes:

- Database I/Os - DB read and DB write I/Os (~34k random I/O)
- Log write I/Os (4k sequential I/O)
- Database maintenance read I/Os (BDM) (256k sequential I/O)
- Log replication read I/Os (~170k sequential I/O)

Note It is important to point out that the performance achieved is well within Microsoft guidelines for Exchange 2010 subsystem performance, and achieved the target IOPS for each mailbox server with extra headroom. For more details about Exchange 2010 I/Os, see "Understanding Database and Log Performance Factors" at http://technet.microsoft.com/en-us/library/ee832791.aspx.

For more details on mailbox server storage performance and thresholds, review the "Mailbox Server Counters" topic: http://technet.microsoft.com/en-us/library/ff367871.aspx.

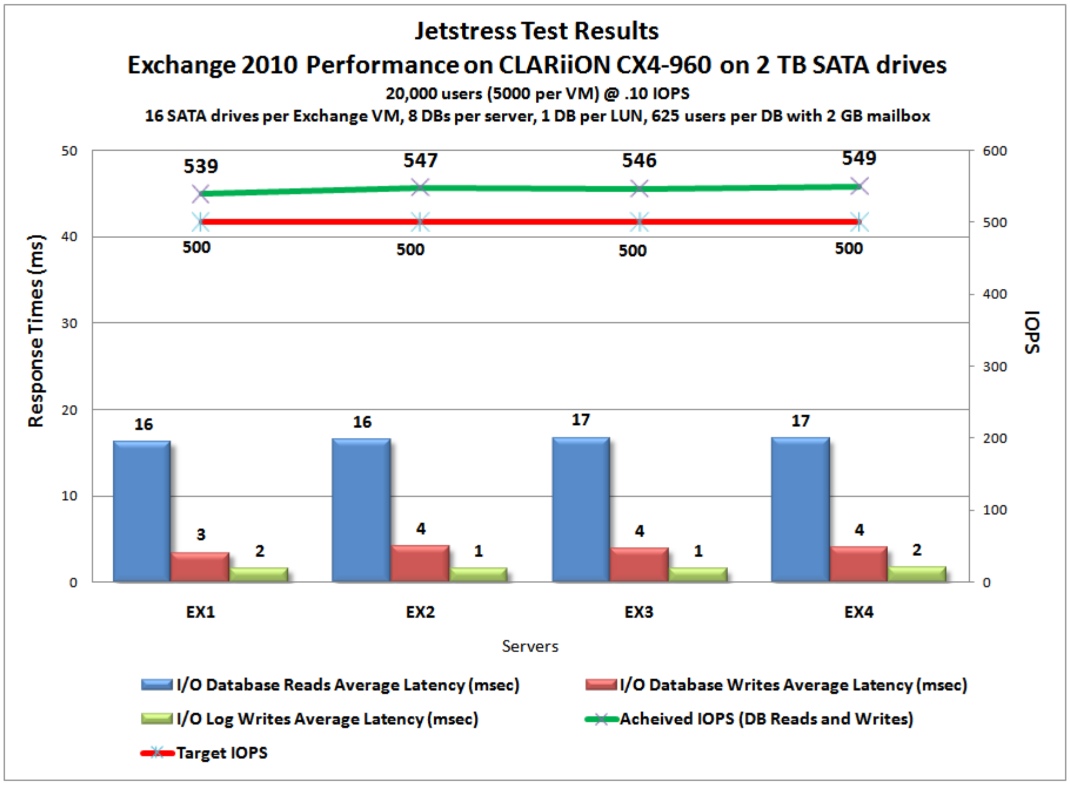

Figure 8. Exchange 2010 Jetstress test summary

Table 19 details performance results for each simulated mailbox server.

Table 19. Jetstress test - individual server details and IOPS summary

Item	Target value	Server 1	Server 2	Server 3	Server 4	Total
Achieved Transactional I/Os /sec (DB reads + DB writes)	500 IOPS per server 2,000 IOPS for four servers	**539**	**547**	**546**	**549**	**2,181**
Database I/Os Reads/sec	N/A	335	341	340	342	**1,357**
Database I/Os Writes/sec	N/A	204	207	206	208	**824**
Database Maintenance I/Os Reads/sec	N/A	253	253	253	253	**1,012**
Log I/Os Writes/sec	N/A	192	196	195	196	**779**
Log Replication I/Os Reads/sec	N/A	7	7	7	7	**28**
Total I/Os		**991**	**1,004**	**1,002**	**1,006**	**4,003**
I/O database Reads average latency (ms)	<20 ms	16	16	17	17	**16**
I/O database Writes average latency (ms)	<20 ms	3	4	4	4	**4**
I/O log Writes average latency (ms)	<10 ms	2	1	1	2	**2**

CX4-960 Storage Array total throughput analysis

Figure 9 shows the total back-end I/O throughput for 20,000 users from four Exchange servers with 5,000 users each at a 0.10 IOPS profile, while Jetstress I/O activity is generated against the CLARiiON CX4-960 array. Results show the array handling 4,000 IOPS (2,000 per SP). See Table 19 for more details.

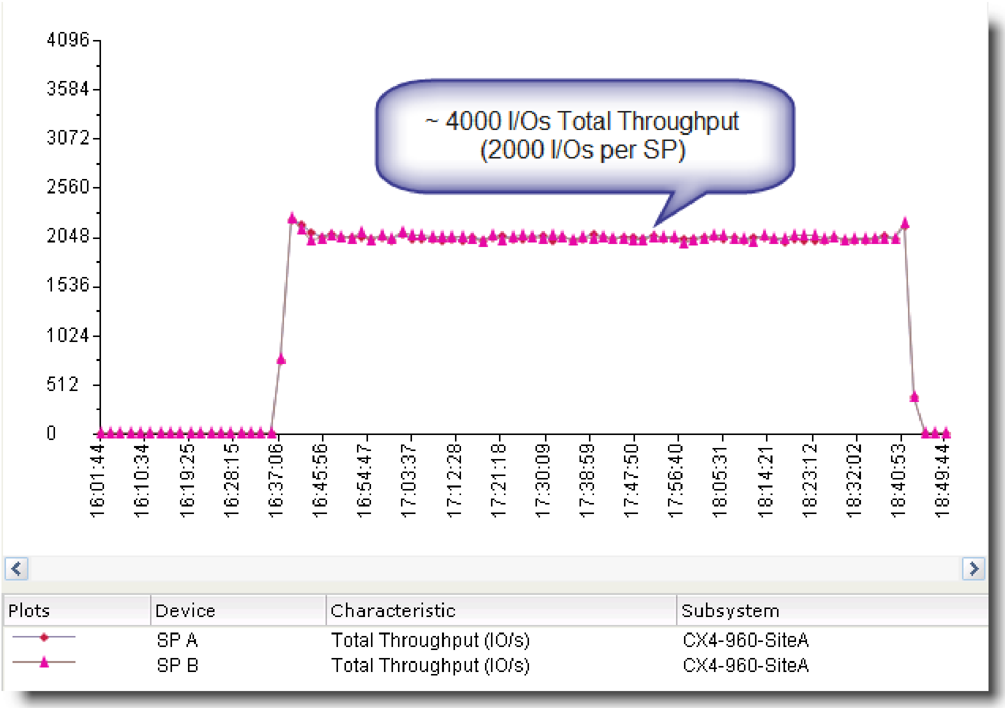

Figure 9. Exchange 2010 Jetstress test summary

Performance of Exchange 2010 storage on SATA drives

Further analysis using EMC CLARiiON Unisphere™ Manager shows SATA drives producing an average 55 random IOPS when the load is mixed with Exchange 2010 large random 34k DB read and write I/Os, and large sequential 256k BDM I/Os. This disk performance is achieved with a RAID 1/0 SATA storage pool utilization at about 65 percent as shown in Figure 10 and Figure 11.

Note Although BDM is considered to be more sequential from the application perspective, it becomes random when mixed at the array back end.

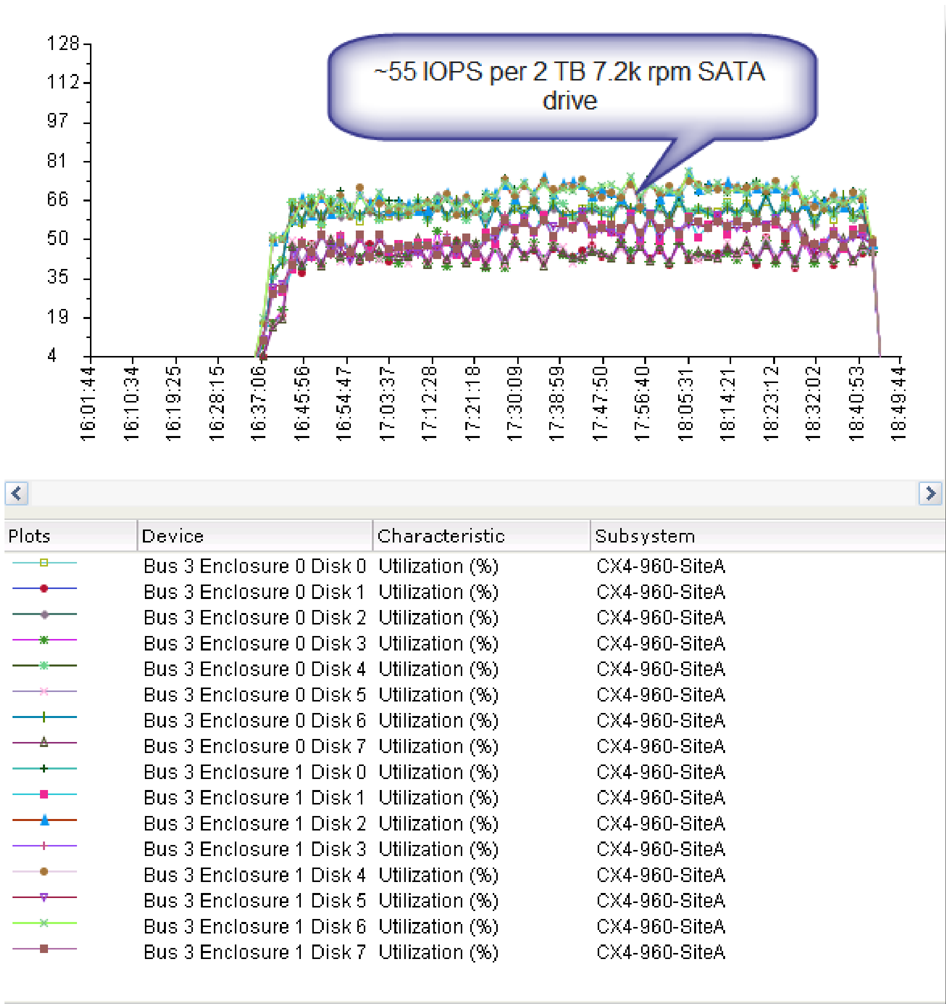

Figure 10. **SATA disk performance**

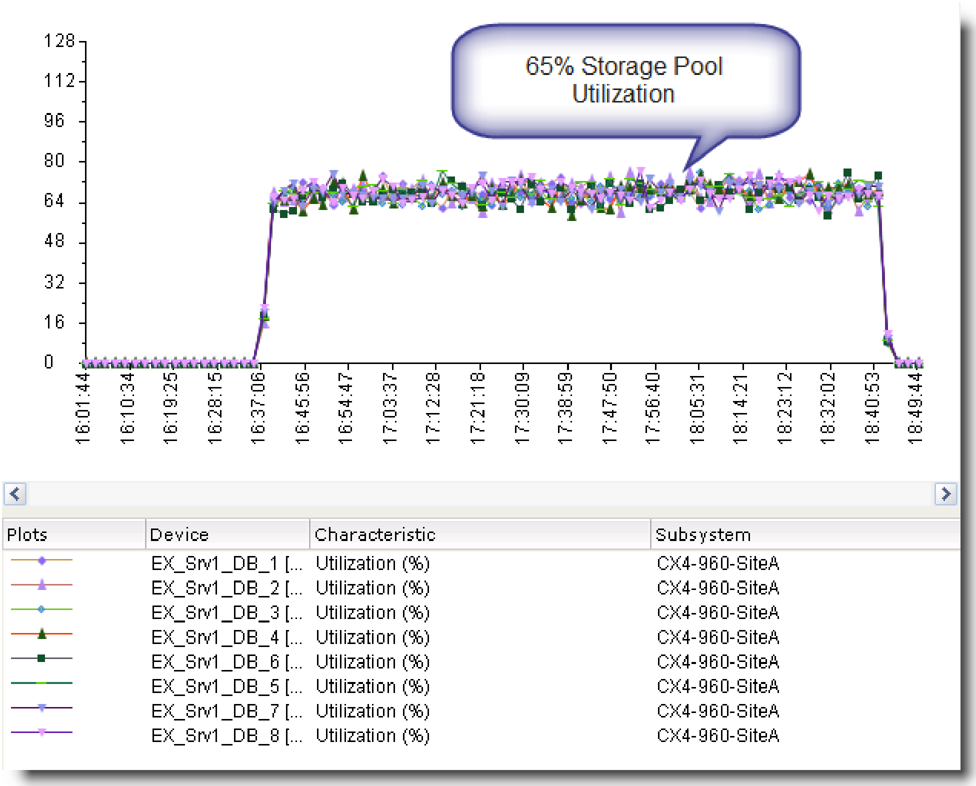

Figure 11. SATA storage pool utilization

Note For additional details about CX4-960 performance with Exchange 2010, review the *EMC CLARiiON CX4-960 60,000 Mailbox Exchange 2010 Mailbox Resiliency Storage Solution* document published on the Microsoft Exchange 2010 Solution Reviewed Program (ESRP) website: http://technet.microsoft.com/en-us/exchange/ff182054.aspx or http://www.emc.com/collateral/hardware/technical-documentation/h7202-esrp-clariion-60000-user-exchange-2010.pdf.