White Paper

MICROSOFT EXCHANGE SERVER BEST PRACTICES AND DESIGN GUIDELINES FOR EMC STORAGE

EMC VNX Series, EMC Symmetrix VMAX systems

- Design best practices
- Sizing best practices
- Exchange Server performance acceleration with EMC flash technologies
- Disaster recovery and high availability solutions

EMC Solutions

Abstract

This white paper provides best practices and key decision points for planning and deploying Microsoft Exchange Server 2010 and 2013 with the EMC® VNX® and EMC Symmetrix® VMAX™ series storage based on EMC Proven Solutions and validated designs.

October 2013

EMC² PROVEN SOLUTIONS

Microsoft Gold Partner

EMC²

Table of contents

EMC²

Microsoft Exchange Server   4
Best Practices and Design Guidelines for EMC Storage:
EMC VNX Series, EMC Symmetrix VMAX systems

Executive summary

In the planning and design phases of a Microsoft Exchange Server implementation, it is important to understand how the application interacts with the storage platform. It is also critical to know which practices to follow in order to avoid problems, achieve the best possible performance, and provide a great user experience.

From an application architecture perspective, Exchange 2013 continues to use the database availability group (DAG) concept with Windows Server failover clustering that was introduced in Exchange 2010 to maintain database high availability. There are now only two server roles: mailbox server and client access server. The mailbox server handles all activity for the active mailboxes on that server. The client access server provides authentication, limited redirection, and proxy services. The client access server offers all the usual client access protocols: HTTP, POP, IMAP, and SMTP.

From a storage architecture perspective, Exchange 2013 offers some significant changes and improvements for storing service, database schema, and search operations.

- The Exchange Store service (Information Store process) has been re-written in managed code (Managed Store) and now each database runs under its own process, which isolates store issues to a single database.

- The Managed Store continues to use Extensible Storage Engine (ESE) as the database engine and works with the Microsoft Exchange Replication service to manage mailbox databases.

- The newly optimized and enhanced mailbox database schema results in reduced user IOPS.

- The Microsoft Exchange Replication service is responsible for all service availability related to mailbox servers.

- The new Managed Store is now integrated with the Search Foundation search engine (also used by SharePoint 2013) to provide more robust indexing and searching.

These changes allow for more storage design options and the use of larger-capacity, slow-rotating disk drives for Exchange data. In light of these new changes, EMC has updated its recommended best practices for Exchange storage design to reflect the new features of Exchange Server 2013.

Purpose of this paper	This paper presents the current set of best practices recommended by EMC for storage design in support of Microsoft Exchange Server 2010 and 2013. It presents guidelines within the context of deploying Exchange on the EMC VNX family of unified storage and EMC Symmetrix VMAX series storage. The paper includes guidelines for deploying Exchange in both physical and virtual environments.
Audience	This white paper is intended for customers, EMC partners, and service personnel considering an implementation of a new messaging environment with Microsoft Exchange Server or considering upgrading from an earlier version of Exchange. We assume that you are familiar with Microsoft Exchange Server, VNX or VMAX series storage, and VMware vSphere or Microsoft Hyper-V virtual environments.
Scope	This document presents storage design best practices recommended by EMC for hosting Microsoft Exchange Server on VNX or VMAX storage in physical or virtual environments. The paper includes sizing and design examples based on EMC's proven building-block approach. Detailed end-to-end implementation instructions are beyond the scope of this document.
Terminology	Table 1 describes several terms used in this document.

Table 1.   Terminology

Term	Description
Background database maintenance (BDM)	The process of Exchange Server 2010/2013 database maintenance that includes online defragmentation and online database scanning. This process produces a large sequential 256 KB I/O read that runs, by default, 24x7 on both the active and passive database copies and can be scheduled to run against active database availability group (DAG) copies. The process cannot be scheduled, throttled, or disabled on passive DAG copies.
Exchange database availability group (DAG)	The base component of the high availability and site resilience framework built into Microsoft Exchange Server 2010/2013. DAG is a set of up to 16 Exchange mailbox servers that provide automatic database-level recovery from a database, server, or network failure.
FAST™ VP	Fully Automated Storage Tiering with Virtual Pools (FAST VP) supports sub-LUN automated tiering for file systems. The feature is available on VMAX and VNX storage systems.
FAST Cache	A VNX storage software feature that enables customers to add various flash drive capacities in order to extend existing cache capacity for better system-wide performance. FAST Cache is available with increased capacity configurations using 100 GB or 200 GB flash drives.
Hypervisor	The layer of software that exists above the hardware and below the management operating system. It creates partitions to provide isolated execution environments and manages each partition's access to hardware resources.
Pass-through disk	A disk that is configured for Hyper-V to "bypass" the host's file system and access the disk directly. This disk can be a physical hard disk on the host or a logical unit on a storage area network (SAN). Pass-through disks do not provide certain virtual hard disk (VHD) features, including VHD snapshots, dynamically expanding VHDs, and differencing VHDs. However, they provide the fastest performance of any VHD type.
Pass-through disk	An option in the Windows Hyper-V server virtualization environment that enables a storage logical unit number (LUN) to directly connect to a virtual machine from the storage area network (SAN).

Term	Description
Raw Device Mapping (RDM)	An option in the VMware vSphere server virtualization environment that enables a storage logical unit number (LUN) to be directly connected to a virtual machine from the storage area network (SAN).
Redundant array of independent disks (RAID)	A method for storing information where the data is stored on multiple disk drives to increase performance and storage capacity and to provide redundancy and fault tolerance.
Server Message Block (SMB)	The protocol used to access file shares across the network.
Storage pool	Virtual constructs that enable data to move dynamically across different tiers of drives (from high performance to lower cost/high capacity, and vice versa) according to the data's business activity. With VNX and VMAX systems, storage pools are fully automated and self-managing.
Thick LUN	A type of pool LUN in which the physical space allocated is equal to the user capacity seen by the host server
Thin LUN	A type of pool LUN in which the physical space allocated can be less than the user capacity seen by the host server.
Virtual hard disk (VHD/VHDX)	The file format for a virtual hard disk, which is the storage medium for a virtual machine on Microsoft Hyper-V. It can reside on any storage topology that the management operating system can access, including external devices, SANs, and network-attached storage. VHD is the default DISK format on Windows Server 2008 R2 and VHDX is the new default format on Windows Server 2012. VHD supports for up to 2 TB disk capacity and VHDX up to 64 TB.
Virtual Machine Disk (VMDK)	The file format for a virtual hard disk, which is the storage medium for a virtual machine on VMware vSphere. VMDK supports up to 62 TB (vSphere 5.5).
XtremSF™	EMC XtremSF is a PCIe flash card deployed in the server to dramatically improve application performance by reducing latency and accelerating throughput. XtremSF can be used as a local storage device to accelerate read and write performance, or it can be used in conjunction with server flash caching software, EMC XtremSW Cache, for accelerated read performance with data protection.
XtremSW™ Cache	EMC XtremSW Cache is intelligent caching software that uses server-based flash technology and write-through caching for accelerated application performance with data protection.

EMC²

Microsoft Exchange Server   8
Best Practices and Design Guidelines for EMC Storage:
EMC VNX Series, EMC Symmetrix VMAX systems

Featured EMC storage

EMC VNX family   The EMC VNX family delivers industry-leading innovation and enterprise capabilities for file, block, and object storage in a scalable, easy-to-use solution. The Next-Generation VNX series enabled with multicore optimization (MCx™) unleashes the power of flash to address the high-performance, low-latency requirements of virtualized applications. This next-generation storage platform combines powerful and flexible hardware with advanced efficiency, management, and protection software to meet the demanding needs of today's enterprises.

When combined with the FAST™ Suite and the new block deduplication technology, VNX delivers the highest performance at the lowest possible midrange cost. VNX also dramatically increases data availability within and across data centers, and simplifies management for storage, application, and virtualization administrators.

VNX Systems are available in a choice ranging from affordable entry-level solutions to high-performance, petabyte-capacity configurations that service the most demanding application requirements. The VNX series is designed to meet the high-performance, high-scalability requirements of midsize and large enterprises. The References section on page 45 provides more information.

VNX Series Hardware

The Next-Generation VNX series hardware includes storage processors for block access and X-Blades and control stations for file access. Table 2 lists the VNX hardware capabilities:

Table 2.   Next Generation VNX Series hardware capabilities

Capability	VNX5200	VNX5400	VNX5600	VNX5800	VNX7600	VNX8000
Maximum drives	125	250	500	750	1000	1500[1]
Supported drive types	Flash, SAS, NL-SAS					
Max usable FAST Cache	600 GB	1 TB	2 TB	3 TB	4.2 TB	4.2 TB
File: # of X-Blades	1 or 2	1 or 2	1 or 2	2 or 3	2 to 4	2 to 8
File: Memory per X-Blade	6 GB	6 GB	12 GB	12 GB	24 GB	24 GB
File: protocols	Common Internet File System (CIFS), NFS, Parallel Network File System (pNFS)					
Block: # of Storage Processors (SP)	2					
Block: Memory per SP	16 GB	16 GB	24 GB	32 GB	64 GB	128 GB
Block: # of cores per SP	4	4	4	6	8	16
Block: protocols	Fibre Channel (FC), iSCSI, Fibre Channel over Ethernet (FCoE)					

---

[1] At the time of publication, only 1,000 drives are supported. Support for 1,500 drives will be available in the first half of 2014.

VNX Series Software

Table 3 lists the software suites that are available for the VNX series.

Table 3.   VNX Series suites

VNX Series Software	Contents
Unisphere® Management Suite (required)	Unisphere, Unisphere Remote, Unisphere Analyzer, Unisphere QoS, VNX Monitoring and Reporting (storage-only version of Watch4Net)
Operating Environment (required)	All protocols, thin provisioning, compression, deduplication, SAN Copy, and ODX Enabler
EMC Storage Analytics for VNX	VNX version of VMware vCenter Operations Manager, EMC Adapter for VNX
FAST Suite	FAST Cache and FAST VP
Security and Compliance Suite	VNX Host Encryption, File Level Retention, VNX Event Enabler (VEE)—Common Anti-Virus Agent and Common Event Publishing Agent
Local Protection Suite	SnapSure™, SnapView™, VNX Snapshots, EMC RecoverPoint® Solutions Enabler (SE) continuous data protection (CDP)
Remote Protection Suite	Replicator, MirrorView™ A/S, EMC RecoverPoint SE continuous remote replication (CRR)
Application Protection Suite	AppSync®, Replication Manager

Advanced data management suites

In addition to the VNX software suites, there are two software packs available. The software packs bundle EMC's advanced data management suites:

- **Total Protection Pack**—Includes Local Protection Suite, Remote Protection Suite, and the Application Protection Suite

- **Total Efficiency Pack**—Includes FAST Suite, Security and Compliance Suite, Local Protection Suite, Remote Protection Suite, and the Application Protection Suite

EMC Symmetrix VMAX series

The EMC Symmetrix Virtual Matrix Architecture (VMAX) series offers high-end storage for the data center. The system scales to a massive 4 PB and consolidates more workloads with a much smaller footprint than other arrays. VMAX seamlessly scales performance, capacity, and connectivity on demand to meet all application requirements. The system supports flash, Fibre Channel, and SATA drives and optimized automatic tiering with FAST VP. The system also supports virtualized and physical servers, including open systems, mainframe, and IBM hosts.

VMAX Series Hardware

Table 4 lists the available VMAX systems with hardware specifications.

Table 4.   VMAX Series Hardware options

	VMAX 10K	VMAX 20K	VMAX 40K
Maximum drives	1,080	3,200	3,200
Usable capacity	1.5 PB	2 PB	4 PB (3.2″ drives) and 2.8 PB (2.3″ drives)
Connectivity	FC, 1GbE, 10 Gbs FCoE, iSCSI	FC, FICON 1GbE, 10GbE, 10 Gbs FCoE, iSCSI	FC, FICON 1GbE, 10GbE, 10 Gbs FCoE, iSCSI
Front-end ports	64	128	128
Local replication	TimeFinder®, EMC RecoverPoint	TimeFinder, EMC RecoverPoint	TimeFinder, EMC RecoverPoint
Remote replication	SRDF®, EMC RecoverPoint	SRDF, EMC RecoverPoint	SRDF, EMC RecoverPoint

The References section on page 45 provides more information.

VMAX Series Features

Table 5 describes the features and benefits provided in the EMC Symmetrix VMAX series systems.

Table 5.   EMC Symmetrix VMAX series features

Feature	Description
Symmetrix FAST and FAST VP	Automates storage tiering to lower costs and deliver higher service levels
Linear scale-out of storage resources	Consolidates multiple arrays into a single VMAX system
Up to 4 PB usable capacity	Seamlessly scales from 48 to 3,200 drives
One to eight VMAX engine scaling	Consolidates more workloads in a smaller footprint with up to eight highly available VMAX engines
Virtual logical unit number (LUN) technology	Transparently moves data to the correct tiers and RAID types as needed
Virtual provisioning	Simply and efficiently allocates, increases, and reclaims storage
Extended distance protection	Replicates data over extended distances and provides zero-data-loss protection
Information-centric security	Provides built in advanced RSA security technology to keep your data safe, reduce risk, and improve compliance

Exchange storage characteristics

In the continuing Microsoft tradition, Exchange 2013 provides a new architecture, enhanced availability features, and further-optimized storage I/O operations. Exchange Server 2013 includes significant changes compared to the Exchange 2010 version and requires more server resources, CPU, and memory. Additional memory resources are needed to support changes made to the Store database engine to reduce demands for disk I/Os. The rendering of mailbox information and a new search indexing system on the mailbox role requires an increase in CPU requirements. Specifically, the storage schema was improved and modified to accommodate high availability and larger mailboxes with fewer I/Os per second (IOPS) per mailbox.

I/O and bandwidth characteristics

To design an efficient storage configuration, you must understand the I/O and bandwidth characteristics of Exchange Server. In general, Exchange I/O has random access to the database and sequential access to the transactional logs. Table 6 presents the changes in user profile characteristics between three versions of Exchange. These estimated measurements include all database volume I/Os, such as database, content indexing, and NTFS metadata, but do not include log volume I/Os. The estimates are based on an average message size of 75 KB. The table shows that there is a 50 percent reduction in I/Os from Exchange 2007 to Exchange 2010 and an additional 33 percent reduction in I/Os from Exchange 2010 to Exchange 2013.

Table 6.   Exchange user profile comparison

Messages sent/received per day	Exchange Server 2007 user IOPS	Exchange Server 2010 user IOPS for protected database copy (mailbox resiliency)	Exchange Server 2013 user IOPS for protected database copy (mailbox resiliency)
50	0.11	0.05	0.034
100	0.18	0.100	0.067
150	0.32	0.150	0.101
200	0.48	0.200	0.134
250	0.64	0.250	0.168
300	1.60	0.300	0.201
350	2.56	0.350	0.235
400	3.52	0.400	0.268
450	4.48	0.450	0.302
500	5.44	0.500	0.335
Comparison is based on average message size of 75 KB			

Mailbox database I/O read/write ratios

With Exchange Server 2013, the larger database cache decreases the number of reads to the database on disk, causing the reads to shrink as a percentage of total I/Os. Table 7 shows the I/O ratios you can expect for active database copies if you follow the recommended memory guidelines.

Table 7.   Exchange 2013 mailbox database I/O read/write ratios

Messages sent/received per mailbox per day	Stand-alone databases	Databases participating in mailbox resiliency
50	1:1	3:2
100	1:1	3:2
150	1:1	3:2
200	1:1	3:2
250	1:1	3:2
300	2:3	1:1
350	2:3	1:1
400	2:3	1:1
450	2:3	1:1
500	2:3	1:1

For more information about Exchange 2013 memory guidelines, see the Exchange Team Blog article.

It is important to point out that the read/write ratios might be different if you use server side cache technologies for your Exchange deployment. For example, EMC XtremSW™ Cache technologies increase read caching, thus changing the overall application read-to-write ratio going to the back-end storage. Using XtremSW Cache with data deduplication also reduces the number of writes to the back-end storage. XtremSW Cache with Exchange on page 36 provides additional details.

Having more writes as a percentage of the total I/O has specific implications when choosing a redundant array of independent disks (RAID) type that has significant costs associated with writes, such as RAID5 or RAID6.

For more information about selecting the appropriate RAID solution for your Exchange deployment, follow the instructions in RAID type selection on page 18. In addition, review the Exchange 2013 online help information, Best practices for supported storage configurations.

Exchange I/O sizes   The Exchange Mailbox server stores mail for all users in an ESE database. The ESE database is randomly accessed and, depending on the Exchange version, uses a specific page size. I/O coalescing can result in more I/Os to disk. For reliability, and in some cases for performance reasons, the database should be on disks that do not contain transaction logs.

Even though the number of user IOPS decreased for Exchange Server 2013 compared to previous versions of Exchange, the size of the I/O has significantly increased.

Microsoft Exchange Server   13
Best Practices and Design Guidelines for EMC Storage:
EMC VNX Series, EMC Symmetrix VMAX systems

Between Exchange 2007 and Exchange 2013, the size of the user I/O increased by four times, from 8 KB to 32 KB. Thus, even though the amount of I/O activity per user has decreased in each new Exchange version, the size of the I/O that the storage subsystem must handle has grown significantly. In addition, there is a 256 KB background database maintenance (BDM) process I/O that continuously runs on the active and passive database copies. Table 8 provides additional information about database and transactional log I/O sizes for three versions of Exchange.

Table 8.   Exchange I/O size changes

I/O Type	Exchange 2007	Exchange 2010	Exchange 2013
Database I/O	8 KB random write I/O	32 KB random I/O	32 KB random I/O
Background Database Maintenance (BDM) I/O	N/A	256KB Sequential Read I/O	256KB Sequential Read I/O
Log I/O	Varies in size from 512 bytes to the log buffer size (1 MB)	Varies in size from 4 KB to the log buffer size (1 MB)	Varies in size from 4 KB to the log buffer size (1 MB)

Exchange background database maintenance process

One important difference in Exchange Server 2013 from previous versions is the way database maintenance is performed. The background database maintenance (BDM) process includes online defragmentation and online database scanning and produces a large 256 KB sequential read I/O.

Note: The BDM process is enabled by default and runs continuously on both active and passive database copies. You can schedule the BDM process for active copies but you cannot schedule or disable it on passive database copies.

In Exchange 2013, the BDM process consumes considerably less bandwidth and produces fewer I/Os than the Exchange 2010 version. In the Exchange 2010 database availability group (DAG) configuration, the BDM process ran once a week and produced about 30 IOPS per database with a throughput of 7.5 MB/s. In Exchange 2013, this process needs to run only once in four weeks and produces about nine IOPS per database with a throughput of 2.25 MB/s per database copy.[2]

Table 9 displays the BDM characteristics for Exchange 2010 and 2013.

Table 9.   **BDM characteristics**

BDM characteristic	Exchange 2010	Exchange 2013
Read I/O size	256 KB	256 KB
Max database scan completion time	1 week	4 weeks
Throughput per database (I/O)	30[2]	9[2]
Bandwidth per database (MB/s)	7.5 MB/s[2]	2.25 MB/s[2]

---

[2] BDM throughput and bandwidth values are estimated based on EMC validation with Jetstress 2010 and Jetstress 2013.

For more details about the Exchange BDM process, see the Exchange Team Blog article on Microsoft TechNet.

Bandwidth considerations for BDM

Although BDM bandwidth for Exchange 2013 is significantly less than the Exchange 2010 version, you should still consider the additional I/O and throughput requirements for the BDM process as part of a storage design that deploys a large number of databases. The effect of BDM on throughput (MB/s) can be significant when hundreds of databases are constrained to a single bus on a storage array. You must ensure sufficient bus-throughput capacity to permit BDM to run against all databases—both active and passive—simultaneously, without affecting performance.

Although BDM I/O is considered sequential to a single database, it will become random when hundreds of databases are deployed on the RAID protected SAN storage. Planning for this requirement will help you to avoid performance issues in the deployment.

Thus, it is still necessary to consider the effect on BDM when designing storage for Exchange Server 2013. When you are planning to deploy a large Exchange configuration on a SAN infrastructure, consider deploying fewer large databases (larger than 500 GB) rather than a large number of small databases (up to 500 GB) to minimize bandwidth requirements.

The Microsoft Exchange 2013 Server Role Requirements Calculator and the EMC Exchange 2010-2013 Designer factor in the BDM requirements. They provide the estimated IOPS and MB/s numbers that you should use to validate your back-end storage to ensure that it can support these requirements and that no problems exist with your design.

Log space considerations

When configuring disks for Exchange, database LUNs generally receive the most attention because there is a misconception that database LUNs alone pose the highest risk for performance bottlenecks. However, database transactions are gated by the completion of their associated log writes. Therefore, you should give as much consideration to log LUNs as to database LUNs during Exchange disk configuration.

You can estimate the storage capacity requirements for a log LUN by considering the following factors:

- **Number of log files generated per mailbox per day**—This number depends on the average number of emails a user sends and receives per day and the average mailbox size.

- **Number of days' worth of logs to maintain**—This number depends on how often the log is truncated. Exchange backups, snaps, or clones, which are typically run nightly, truncate log files. Differential backups, snaps, and clones do not truncate the log files. If the backup/snap/clone schedule includes only weekly full and daily differential backups/snaps/clones, the log LUN space needs to be larger than an entire week of log file space to accommodate both backup and replay data during a restore operation.

- **Storage required for mailbox moves**—Many large companies move a percentage of their user mailboxes nightly or weekly to different databases, servers, or sites. When a mailbox is moved from one database to another,

transaction logs are generated on the target log LUN and are roughly the same size as those generated on the source.

- **Log growth overhead factor**—For most deployments, EMC recommends that you add an overhead space of 20 percent to the log size when you create the log LUNs (after all of the other factors have been considered) to ensure that the required space is available in the event of unexpected log generation.

- **High availability factors**—The implementation of high availability features, such as DAG, increases storage requirements.

Exchange 2013 transaction logs require the same amount of space as Exchange 2010 and you can compute it using the same method. To perform sizing for log space, first determine the base guidelines for the number of transaction logs generated per-user, per-day, using Table 10. As in previous version, Exchange 2013 log files are 1MB in size, making the calculations for log capacity very easy and straightforward. Table 10 provides references for logs generated for each user profile. The Microsoft Exchange Team blog article and Appendix A: Exchange manual storage sizing process provide details for the sizing process.

Table 10.   Number of transaction logs generated per mailbox profile

Message profile (75 KB average message size)	Number of transaction logs generated per day
50	10
100	20
150	30
200	40
250	50
300	60
350	70
400	80
450	90
500	100

Exchange storage design best practices

Storage design is one of the most important elements of a successful Microsoft Exchange Server deployment. To achieve a storage design that is optimal for reliability, performance, cost, and ease of use, it is critical that you follow the recommended guidelines provided by the storage vendor.

This section provides general best practices for deploying Exchange Server 2010 and 2013 on the EMC mid-range VNX and Enterprise VMAX storage systems, including recommendations for using specific array features.

Since the virtualization of an Exchange environment requires its own set of considerations, this section also includes guidance on this subject.

Disk type selection   One of the first key decisions you must make when designing Exchange Server storage is to select the type or types of disks that best match your requirements. The types of disks that are appropriate for your Exchange Server deployment depend on a number of factors, including mailbox size limits and IOPS requirements.

EMC currently offers the following disk types with VNX and VMAX series storage:

- **SATA and NL SAS (7.2k rpm)**—SATA disks (available on VMAX) and NL SAS disks (available on VNX) are a good fit for the less demanding I/O requirements of Exchange Server 2010/2013. These disks support large mailboxes at a relatively low cost. Large capacity SATA or NL SAS disks are typically the best choice for larger mailbox sizes and average to heavy I/O user profiles.

- **SAS (10/15k rpm)**—SAS disks (an improvement over SCSI drives) provide high capacity with moderate I/O speed, which makes them highly suitable for Exchange Server 2010/2013 environments with high IOPS per user requirements.

- **FC (10/15k rpm)**—Fibre Channel disk drives provide high read/write speeds. These disks are ideal for high I/O requirements but might not be appropriate for high capacity requirements.

- **Flash** — EMC is a pioneer in flash drive technology. Flash drives have the highest I/O speed with low power consumption. In general, flash drives might not be appropriate for Exchange data because of their small capacity, but they are a good choice when used as FAST Cache, or as an additional tier in a FAST VP storage pool to handle any unanticipated I/O spikes.

Many IT organizations are considering deploying larger mailboxes for their customers. Since the Exchange Server 2010/2013 storage engine has a lower I/O requirement, low-cost NL SAS and SATA drives are ideal disk types for most environments and are, by far, the most popular choice.

For environments with very high I/O requirements but a moderate mailbox size, 10K rpm or 15K rpm FC or SAS drives are an appropriate choice, since these drives can handle high I/O requirements more efficiently than SATA or NL SAS disks.

Follow these general guidelines for selecting appropriate disk type for Exchange data:

- For lower IOPS requirements and higher mailbox capacity requirements, use SATA or NL SAS disks.

- For higher IOPS requirements and lower mailbox capacity requirements, use large capacity FC or SAS disks.

Exchange IOPS per disk

Different disk types generate different numbers of Exchange Server 2013 IOPS. Consider this when calculating the disk requirements for your environment. Table 11 provides Exchange server IOPS per disk values from the most recent Exchange validation with Jetstress on EMC VNX and VMAX storage. The numbers are based on large 32 KB Exchange database random I/O workloads and small sequential log I/Os for all RAID types.

Important! EMC strongly recommends using the values from Table 11 when calculating the Exchange 2010/2013 disk requirements for deployment on VNX and VMAX storage arrays. These values are also valid for the EMC Exchange Designer.

Table 11.   Exchange Server 2010/2013 IOPS for various disk types on EMC storage

Disk type	Exchange Server 2010/2013 database IOPS per disk (random workload)		Exchange Server 2010/2013 database logs IOPS per disk (sequential workload)
VNX/VMAX	VNX	VMAX	VNX and VMAX
7.2 K rpm NL SAS/SATA	65	60	180
10 K rpm SAS/FC	135	130	270
15 K rpm SAS/FC	180	180	450
Flash	1250	1250	2000

RAID type selection

Selecting an appropriate RAID type for the storage environment is another important decision point for a successful implementation of Exchange Server. You can use any RAID type, provided there are enough disks to handle the I/O and storage capacity requirements. In order to select an appropriate RAID type for the environment, consider the specific customer's performance, capacity, and high availability requirements.

VNX and VMAX storage support most common RAID types used for Exchange storage configuration, such as RAID 1/0, RAID 5, and RAID 6. Each RAID type provides different performance, capacity, and protection levels.

- **RAID 1/0** provides data protection by mirroring data onto another disk. This produces better performance and minimal or no performance impact in the event of disk failure. RAID 1/0 provides protection from a single disk failure in a RAID set. In general, RAID 1/0 is the best choice for Exchange Server environments with high I/O requirement and small to medium mailbox deployment.

Microsoft Exchange Server
Best Practices and Design Guidelines for EMC Storage:
EMC VNX Series, EMC Symmetrix VMAX systems

18

- **RAID 5** data is striped across disks in large stripe sizes. The parity information is stored across all disks so that data can be reconstructed. This can protect against a single-disk failure. For parity protection, RAID 5 generates one parity bit for every write. RAID 5 is most appropriate in Exchange Server environments with low I/O requirements and large mailbox deployment.

- **RAID 6** data is also striped across disks in large stripe sizes. However, two sets of parity information are stored across all disks so that data can be reconstructed if required. RAID 6 can accommodate the simultaneous failure of two disks without data loss. RAID 6 generates two parity bits for every write. RAID 6 is most appropriate in Exchange Server environments with low I/O requirements and large mailbox deployment.

Table 12 compares RAID types for performance, storage utilization and protection levels.

Note: The write penalty value becomes important when performing I/O calculations for disk requirements. Appendix A: Exchange manual storage sizing process provides calculation details.

Table 12.   RAID type comparison

RAID type	Write Penalty	Performance	Storage utilization	Protection Level
RAID 1/0 striping + mirroring	2	High	Low	Medium
RAID 5 striping + parity	4	Medium	High	Medium
RAID 6 striping + double parity	6	Low	Medium	High

General best practices for Exchange Server on EMC storage

To achieve the best storage performance with Exchange Server 2010/2013 on EMC VNX and VMAX storage, follow these general storage design best practices:

Sizing

- Use the Exchange storage building block design approach whenever possible to scale your configuration. Exchange storage design provides more information.

- Always calculate for performance, capacity, and bandwidth requirements.

- Select the appropriate disk and RAID type that meets both the performance and capacity requirements.

- Although the BDM throughput and bandwidth requirements are reduced in Exchange 2013 compared to Exchange 2010, to minimize BDM impact, deploy fewer large databases (500 GB to 2 TB), rather than many small databases (under 500 GB).

- Although database sizes larger than 2 TB are supported (up to 16 TB), ensure that the hypervisor, replication, and backup software support LUNs that are larger than 2 TB. Also, make sure to account for a 20 percent free LUN space buffer and any other data that resides on the same volume as the database when calculating your storage requirements. For example, a 2 TB database cannot fit on a 2 TB LUN.

- When using Exchange DAGs, the recommended maximum database size is 2 TB. Without DAGs, the recommended maximum database size is 200 GB. Plan a database size that is appropriate for your specific requirements. Consider backup and restore times when choosing a database size.

- For best performance and predictable service levels, isolate the Microsoft Exchange Server workload (database and logs) from other I/O-intensive applications and workloads by assigning the Exchange workload to its own set of disks. This simplifies troubleshooting in the event of storage-related performance issues. The single exception to this guideline is the use of properly sized Symmetrix FAST VP.

Database replication

- If your storage design does *not* include DAG replication, you can place the database files from one database and the log files from another database on the same physical disks. However, avoid using the same physical disks for both database and log files for the *same* database. In this case, neither the database files nor the log files are adequately protected against disk failure.

  Note: This recommendation does not apply if storage design employs JBOD deployments with multiple DAG copies, where one physical disk can be used for storing database and log files for the same database.

- If your storage design includes DAG replication, you can place both the database and log files from the same database on the same disks, since there is a copy of the database and log files on a different set of disks. Deploy each

DAG copy on its own set of physical disks. For example, do not put multiple DAG copies of the same database into the same RAID group or storage pool.

Mailbox server

- Format Windows NTFS (or ReFS on Windows Server 2012) volumes used for Exchange databases and logs with an allocation unit size of 64 KB.

- If using ReFS on Windows Server 2012, be sure to install the Microsoft hotfix to correct the problem with heavy defragmentation of Exchange databases, as described in article KB2853418.

- Install the Microsoft hotfix KB2870270 on the Windows Server 2012 hosts in your environment.

- Install EMC PowerPath® for optimal path management and maximum I/O throughput. In virtual deployments, install PowerPath on the hypervisor host.

- Use a minimum of two host bus adapters (HBAs) per server with each HBA connected to at least two array front-end/director ports to avoid a single point of failure.

Storage Array

Spread the load as evenly as possible across storage array resources (that is, VMAX engines, VNX storage processors, buses, and so on).

Virtual storage provisioning with Exchange

Exchange administrator challenges

One of the biggest challenges Exchange administrators face is balancing the amount of storage space to provision in the datacenter to support mailbox quotas for all users. Administrators typically must allocate all mailbox space up front based on anticipated storage growth. They do this to reduce the management costs and application downtime incurred when they need to add more storage as their messaging environment grows. It is likely that most users will not fill their mailbox quotas in the next 3 to 5 years, so all fully provisioned physical storage capacity will be underutilized. This generally results in the overprovisioning of Exchange storage capacity, which then leads to higher costs and an increase in power, cooling, and floor space requirements, and lower capacity utilization rates. Even with careful planning, it might be necessary to provision additional storage in the future, which could require Exchange downtime.

During Exchange migrations from older versions, a typical user mailbox is much smaller than what the Exchange administrator provisions on the new system. For example, a user who is migrated from a 250 MB mailbox quota on the old server to a 10 GB mailbox quota on the new server will most likely fill the new mailbox in the next 3 to 5 years.

Virtual provisioning technology addresses these concerns. Using virtual provisioning, Exchange administrators can purchase and provision only the amount of storage necessary to accommodate the first one to two years of the solution—perhaps 2 GB—and then add more storage at a lower cost as requirements change. Administrators can implement this solution without any interruption or downtime for the end users, because the storage capacity presented to the Exchange servers is based on their fully provisioned mailbox quotas, which is 10 GB in this example.

Virtual provisioning

Virtual provisioning, generally known in the industry as "thin provisioning," enables organizations to reduce storage costs by increasing capacity utilization for Exchange Server environments. Virtual provisioning presents an application with more capacity than is physically allocated. The physical storage is allocated to the application as needed from a shared pool. Virtual provisioning also helps to reduce power, cooling requirements, and capital expenditures. EMC VMAX and VNX systems offer virtual provisioning technology that can simplify storage management and reduce application downtime.

Unlike traditional LUNs, where physical storage capacity is fully allocated during LUN creation, thin LUNs typically have less physical storage allocated to them during LUN creation. For thick LUNs, the physical space allocated is equal to the user capacity seen by the host server. For thin LUNs, the physical space allocated can be less than the capacity seen by the host server. The physical storage that is used to supply disk space to thin LUNs comes from a shared storage pool. The pool is comprised of physical disks that provide the actual physical storage to support the thin LUN allocations. The array-based mapping service builds and maintains all of the storage structures based on minimal high-level user input. Physical storage is automatically allocated only when new data blocks are written. From a performance perspective, thick and thin LUNs can provide all the necessary requirements for any Exchange workload. The References section on page 45 provides more information.

Automated tiering for Exchange data

EMC FAST VP technology automatically optimizes performance in a tiered environment reducing overall costs, footprint, and management efforts. FAST VP delivers automation and efficiency to the virtual data center and offers the following cost and performance benefits:

- FAST VP maximizes utilization of flash drive capacity for high IOPS workloads and maximizes utilization of low IOPS larger capacity NL SAS/SATA drives for capacity-intensive applications like Exchange.

- You can set policies to automatically tier data based on I/O, avoiding the pre-provisioning tasks of determining the tiers on which to assign data.

- You can choose to have data placed on the highest or lowest available tier, ensuring that performance and cost commitments are met.

- You can define data movement schedules to minimize FAST management responsibilities.

Whether to use FAST VP with Exchange Server on VNX or VMAX storage arrays depends entirely on customer requirements. If the customer needs FAST VP for all applications or needs to be able to handle unanticipated spikes in performance demand, FAST VP is a good fit. FAST VP reduces administration workloads and provides faster space and I/O issue resolution.

Changes in the Exchange 2010 and 2013 storage architecture has resulted in lower I/O to storage devices as well as the trend to deploy larger mailboxes, so that many Exchange designs are capable of utilizing high capacity low RPM disks (for example, 7.2 k rpm SATA or NL SAS). Nevertheless, there are Exchange configurations with considerably higher I/O demands and smaller mailbox capacity requirements that would benefit from adding additional faster drive tiers into an existing storage pool created with slower NL-SAS or SATA drives. This solution eliminates problems in Exchange application performance and provides an option for customers who want to use larger, less expensive NL-SAS/SATA drives but require the speed of faster SAS/FC or flash disks.

For best practices and recommendations for configuring FAST VP on VNX and VMAX see FAST VP with Exchange on VNX and Symmetrix FAST VP with Exchange on VMAX.

To ensure that FAST VP will benefit your design, evaluate your current Exchange configuration to identify any hot spots. For specific FAST VP configuration guidelines, review the best practices described later in this document, as well as the VNX and VMAX product documentation.

- For EMC VNX, the *EMC FAST VP for Unified Storage Systems white paper* describes Fully Automated Storage Tiering for Virtual Pools (FAST VP) technology, its features, and implementation guidelines.

- For EMC VMAX, the *EMC Symmetrix VMAX with Enginuity white paper* provides more information.

VNX-specific storage design guidelines

The EMC VNX family of unified storage offers various features and capabilities including virtual provisioning, storage tiering with FAST VP, FAST Cache, deduplication and compression, replication and data protection with SnapView (snaps or clones) or VNX Snapshots, and so on. It is important to understand which of these features are compatible with and beneficial for your Exchange Server environment. This section presents answers to some typical questions about the deployment of Exchange Server on VNX storage.

In addition to general best practices already described, consider the following best practices when configuring Exchange on VNX storage.

System memory and page settings

On the new series of VNX systems (VNX5200, VNX5400, VNX5600, VNX5800, VNX7600, VNX8000), system page size and system memory settings are configured automatically and do not need to be set by the user.

For previous VNX system series (VNX5100, VNX5300, VNX5500, VNX5700, VNX7500), follow these recommendations for page size and memory settings:

- Set the page size to 16 KB

- Enable both read and write cache

- Allocate 1 GB of system memory to read cache and the remaining memory to write cache.

Pools or RAID Groups

On VNX systems, you can configure either traditional RAID groups or storage pools for Exchange data. Traditional storage provisioning with RAID groups restricts the number of disks you can have in a group to 16. Storage pools, on the other hand, enable you to manage potentially hundreds of disks at a time. The use of storage pools simplifies storage provisioning. Such pool-based provisioning provides benefits similar to metaLUN striping across many drives but, unlike metaLUNs, storage pools require minimal planning and management efforts. Storage pools support the same RAID protection levels as RAID 5, RAID 6, and RAID 1/0. When using storage pools, you can create multi-tiered pools with different RAID and disk types. Storage pools also allow you to take advantage of advanced data services (like FAST VP, compression, and deduplication) and data protection options, such as VNX Snapshots.

When using Storage Pools on VNX, depending on requirements, the following guidelines apply:

- When using single-tier pool (without FAST VP), use homogeneous storage pools (pools with the same disk type) for Exchange data and logs.

- When using multi-tiered pool (with FAST VP), use heterogeneous storage pools (pools with mixed disk types) for databases, and create a separate single-tier pool for transactional logs.

- Design and grow storage pools by using the appropriate multiplier for best performance:
    - For RAID 1/0: 4+4

- For RAID 5: 4+1 or 8+1
- For RAID 6: 6+2 or 14+2

- To use VNX Snapshots, you must configure Exchange storage (databases and logs) in storage pools.

- When using pools with thin LUNs, be sure to use storage pool utilization capacity threshold alerts to monitor the pools and keep them from running out of space.

When using RAID groups, the following guidelines apply:

- If you are using RAID groups with RAID 5 or RAID6, do not stripe metaLUNs across multiple RAID groups. Doing so can reduce performance.

- Remember that the maximum number of disks allowed in a RAID Group is 16.

Thick or thin LUNs with Exchange

When using storage pools, you can create thick or thin LUNs for the Exchange database and logs. When creating thick LUNs, the physical space allocated is equal to the user capacity seen by the mailbox server. When you create thin LUNs, the physical space allocated can be less than the user capacity seen by the mailbox server. From a performance perspective, thick and thin LUNs can fulfill all the necessary requirements for any Exchange workload on Next-Generation VNX systems. However, thin LUNs might present some additional overhead due to metadata tracking.

Using thin LUNs can provide significant storage savings when deploying large mailboxes because you can create LUNs with the required user capacity but with less physical capacity in the storage array.

Use these guidelines to determine the best LUN type for deploying Exchange on VNX:

- Deploy thin LUNs for moderate user workloads.

- Deploy thick LUNs for heavy user workloads.

- Always perform thin and thick LUNs optimization before Jetstress validation and Exchange deployment. VNX storage pool optimization for Exchange workload provides more information.

- If maximum possible performance without the future need for VNX Snapshots is required and it is acceptable to commit all physical storage from Day 1, use thick LUNs.

- Use thin LUNs if your deployment requires protection with VNX Snapshots.

- When the LUNs (thick or thin) are protected with VNX Snapshots, Enable FAST Cache.

FAST Cache with Exchange

FAST Cache on VNX storage systems enables customers to use flash drives to extend existing storage array cache capacity for better system-wide performance. FAST Cache on VNX systems provides immediate relief for hot spots and boosts performance for small-block random I/O application workloads with high locality and with frequent and unpredictable large increases in I/O activity. The part of an application dataset that is frequently accessed (data re-hits) is copied to the FAST Cache, so the application receives an immediate performance boost for data that resides on FAST Cache-enabled LUNs.

Because of the large-block I/O, lower Exchange 2010/2013 user I/O profile requirements, and very small working dataset (2-5%), there is rarely a need for additional cache boosts for most configurations on VNX systems. Exchange uses a considerable amount of host cache and employs data coalescing, which sometimes works counter to FAST Cache technology. However, the use of FAST Cache on VNX will benefit Exchange performance in certain configurations. Specifically, EMC recommends that you enable FAST Cache on thin and thick LUNs when VNX Snapshots are enabled.

Follow these guidelines when using FAST Cache with Exchange:

- Segregate the database volumes from the log volumes by placing each volume type in a separate storage pool or RAID group.

- Enable FAST Cache only on database LUNs. Do not enable FAST Cache on log LUNs, even if you are using VNX Snapshots. Enabling FAST Cache on log LUNs will not provide any benefit because of the sequential nature of Exchange transactional logs I/O.

  Note: When using storage pools, FAST Cache can be enabled or disabled at the storage pool level only and not at the LUN level.

- When VNX Snapshots are configured for Exchange databases on storage pool LUNs (thick or thin), enable FAST Cache.

- Be aware that FAST Cache is a shared global array resource and can be used for other applications configured on the storage array. As a general sizing recommendation, use a 1,000:1 Exchange data to FAST Cache ratio. That is, for 1,000 GB of Exchange data, provision approximately 1 GB of FAST Cache.

FAST VP with Exchange on VNX

FAST VP uses storage pools to provide sub-LUN tiering, enabling the utilization of multiple tiers of storage simultaneously. FAST VP optimizes the storage system TCO as it moves data to the appropriate storage tier based on sustained data access and demand over time. Compared to FAST Cache, which is a global resource in the array, FAST VP is used on a dedicated storage pool to benefit a particular application. When FAST VP is used in a multi-tiered storage pool with thin LUNs, all metadata is automatically promoted to the highest tier, thus improving the overall performance for Exchange. To ensure that FAST VP will benefit your design, evaluate your current Exchange configuration to identify any hot spots. When designing FAST VP for Exchange on VNX, follow these guidelines:

- When using FAST VP, do not place database files and log files in the same pool. Log files have a low I/O requirement and do not need to be moved to a higher tier.

- When using FAST VP with DAG, never place DAG copies of the same database in the same pool on the same disks.

- When using FAST VP, set the FAST tiering policy for the participating pool-based LUNs to **Start High then Auto-Tier**.

- When sizing the FAST VP tiers, ensure the skew (working set) for the application will fit with room to avoid thrashing between tiers. Exchange's working set is typically around two to five percent.

VNX storage pool optimization for Exchange workload

When a VNX storage pool is created, the process automatically creates private objects that are not visible using traditional reporting tools. Within the context of a pool, private objects are RAID Groups and LUNs created as resources to be used by the mapper service using the pool. Private objects of a pool align with the best practice recommendation for disk count per RAID group, for example, 5 for RAID5 and 8 for RAID6 and RAID 1/0.

VNX pool-based LUNs allocate slices provisioned from the private objects in the pool. Allocation is essentially the same for both thin and thick LUNs, although the logical blocks within the slices map differently to their own visible LUN. When Exchange is deployed on storage pool-based LUNs (thin or thick), the VNX file system evenly provisions and allocates slices from private pool resources to each LUN. Both thin and thick LUNs allocate a minimum number of slices from those private pool resources on creation. For thick LUNs, the space and slices are fully provisioned during LUN creation. For thin LUNs, the slices are allocated on demand as new data is written to the LUN.

When Exchange databases created on thin LUNs are slowly populated with users' email, the VNX file systems try to allocate requested storage evenly to all the LUNs. However, with Exchange Jetstress this behavior is different because of the way Jetstress creates the databases.

When Jetstress creates the databases for pre-deployment storage validation, it creates the first database and then copies it in parallel to the rest, as defined in the configuration. If your design employs configuration with a single database per LUN, you might see that the first database has better performance with lower latencies than the others. This is because more slices are allocated to the first LUN (where the first database resides) than the others. Figure 1 shows how Exchange Jetstress creates databases. In the example, Jetstress creates DB1 and then simultaneously copies it to the other databases.
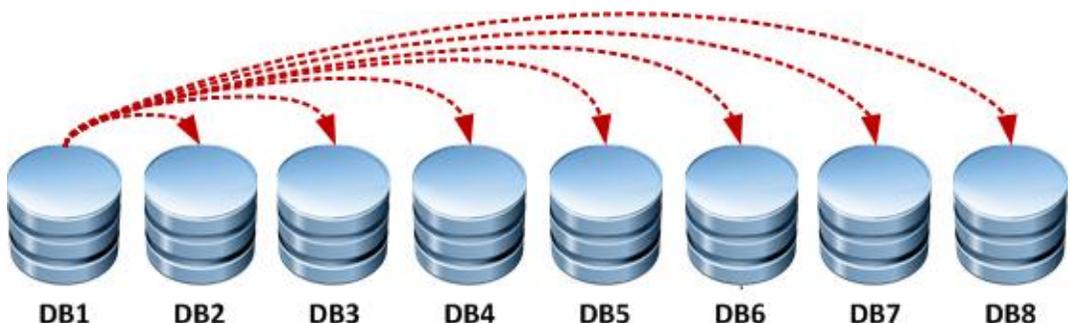
Figure 1.   Exchange Jetstress database initialization process

To ensure that all databases that reside on storage pool-based LUNs perform equally, use the **VNX Pool Optimization utility**. This utility performs optimization of all the LUNs within the storage pool by warming up and pre-allocating the slices for each LUN, allowing each LUN to perform uniformly.

The **ESI for VNX Pool Optimization utility** provisions and optimizes VNX storage pool thick and thin LUNs for maximum performance after the LUN creation and before disk partitioning on the Exchange server. It is a best practice to use this tool when preparing the environment for Jetstress validation to ensure uniform, deterministic high performance and low latencies across all LUNs within a VNX storage pool.

The *EMC Storage Integrator for VNX Pool Optimization Release Notes*[3] available on EMC Online Support provides more details and instructions. EMC Storage Integrator (ESI) for Windows provides additional information about ESI.

---

[3] This document is scheduled for publication in mid-November 2013.

VMAX-specific storage design guidelines

The EMC Symmetrix VMAX family of storage systems offers many advanced features and capabilities that can be used for Exchange workloads, including virtual provisioning, storage tiering with FAST VP, and replication with TimeFinder. It is important to understand which of these features are compatible with and beneficial to your Exchange Server environment. This section presents answers to some typical questions about the deployment of Exchange Server on VMAX storage.

In addition to general best practices already described, consider the following best practices when configuring Exchange on VMAX storage:

- Use Symmetrix Virtual Provisioning™.

- Share database and log volumes across the same disks, but separate them into different LUNs on the same hosts.

- For optimal Exchange performance, use striped meta volumes.

Symmetrix Virtual Provisioning for Exchange

EMC recommends using Symmetrix Virtual Provisioning for Exchange Server 2010 and 2013 workloads. Performance of thin devices and thick devices is equivalent on VMAX and, in most cases, using thin pools can reduce the initial storage requirement.

If you use thin pools, you must ensure that the initial disk configurations can support the I/O requirements. You can configure a thin pool to support a single Exchange building block or multiple building blocks, depending on the customer's requirements.

Consider the following best practices when using Virtual Provisioning on Symmetrix VMAX for Exchange:

- Ensure that the initial disk configurations can support the I/O requirements. You can configure a thin pool to support a single Exchange building block or multiple building blocks, depending on customer requirements.

- Use Unisphere for VMAX to monitor the thin pool utilization and prevent the thin pools from running out of space.

- Install the Microsoft hotfix KB2870270 on the Windows Server 2012 hosts in your environment.

Symmetrix FAST VP with Exchange on VMAX

As previously discussed, because of the reduction in user I/O profiles in Exchange 2012 and 2013 and the trend to deploy larger mailboxes, many Exchange designs are capable of using high capacity and low rpm disks (for example, 7.2 k rpm SATA). To ensure that FAST VP is appropriate for your design, evaluate the current Exchange configuration to identify any hot spots.

When designing FAST VP for Exchange 2010/2013 on VMAX, follow these guidelines:

- Separate databases and logs onto their own volumes, but you can share database and log volumes across the same disks.

- For Exchange workloads, exclude transaction log volumes from the FAST VP policy or pin all the log volumes into the FC or SATA tier.

- Select **Allocate by FAST Policy** to allow FAST VP to use all tiers for new allocations based on the performance and capacity restrictions. This is a new feature introduced in the Enginuity™ 5876 code.

- When using FAST VP with Exchange DAG, do not place DAG copies of the same database in the same pool on the same disks.

EMC²

Virtualization best practices for Exchange Server on EMC storage

Overview

In recent years, virtualizing most mission-critical applications as well as the Microsoft Exchange Server environment has become a popular choice for many companies because of the many benefits virtualization can provide. With the new, more powerful servers, you can achieve high server consolidation ratios, which can yield significant cost savings.

Exchange Server 2010 and Exchange Server 2013 are both supported on hardware virtualization software such as Microsoft Hyper-V, VMware vSphere, and any third-party hypervisor that has been validated under the Windows Server Virtualization Validation Program (SVVP) at http://www.windowsservercatalog.com/svvp.aspx. SVVP also provides information about how EMC storage and software support the virtualization of Exchange Server environments.

For more information about Exchange Server system requirements for hardware virtualization, visit Microsoft TechNet:

- Exchange 2010 Virtualization

- Exchange 2013 Virtualization

The Troubleshooting Microsoft Exchange Server Performance article provides more information about identifying and addressing performance bottlenecks in an Exchange system.

General storage guidelines for Exchange virtualization

Virtualization of the Exchange Server environment generates additional requirements to ensure high-availability and a great user experience. These requirements include hypervisor host sizing, physical infrastructure configuration, and failure domain considerations, as well as Exchange database placements to ensure high availability.

For the most part, when designing for Exchange virtualization you should apply the same core design principles as you would for physical deployment. You must plan for performance, reliability and capacity. Here are some of the design principles you should follow when considering virtualizing a Microsoft Exchange Server:

Hypervisor

- Install PowerPath® for Windows on the Windows Hyper-V server or PowerPath/VE on the vSphere ESXi server for performance optimization, path load balancing and management, and I/O path failure detection.

- Size physical host servers to accommodate the required number of supported guests.

- Disable hypervisor-based auto tuning features, including dynamic memory, storage tuning, and rebalancing.

- Ensure that each host has at least four paths to the storage. To accomplish this, the host requires at least two HBA ports.

- Know your hypervisor's limits, processor limits, and memory limits. For example, establish a maximum of 256 SCSI disks per physical host or cluster.

Storage

- All storage used by an Exchange guest machine for storage of Exchange data must be block-level storage.

- Exchange 2013 virtual disks with Exchange data and logs (VHD/VHDX) can be stored on SMB 3.0 file shares. The shares can be presented to a Hyper-V server from Windows Server 2012 or other implementations of SMB 3.0. The VHD storage must not have direct access to shares from Exchange (for example, no direct UNC path to a share: \\server\share\db1\db1.edb).

  Note: NFS shares or SMB implementations other than SMB 3.0 are not supported. The Exchange storage supportability and guidelines section provides more information.

- Depending on technical requirements, use the appropriate volume types for Exchange databases and logs. On VMware vSphere, use VMFS, RDMs, or iSCSI. On Microsoft Hyper-V, use a SCSI pass-through device, VHD/VHDX, or iSCSI.

- You can configure iSCSI storage to use an iSCSI initiator inside an Exchange guest virtual machine. However, this configuration results in reduced performance if the network stack inside a virtual machine is not full-featured (for example, not all virtual network stacks support jumbo frames).

- For VMware vSphere virtual machines, use virtual LSI Logic SAS (vSCSI) or Paravirtual SCSI (PVSCSI) adapters. To determine which one to use, refer to Virtual SCSI Adapters on page 34.

- Use multiple vSCSI adapters. This helps to avoid inducing queue depth saturation within the guest OS.

- Add all four SCSI controllers to the virtual machine and equally distribute disks across the four available vSCSI adapters

- Place the Exchange storage on disks that are separate from the guest OS physical storage.

Virtual machines

- The number of users supported by each mailbox server virtual machine depends on the hypervisor type, physical server hardware capabilities (such as CPU and memory characteristics), and high availability requirements.

- Size virtual machines according to the Exchange roles.

- Deploy one DAG member per host. If there are multiple DAGs, they can be co-located on the same host. Spread DAG copies across multiple physical servers to minimize potential downtime in the event of physical server issues.

- When calculating guest virtual machine CPU requirements, include additional hypervisor overhead:

  - 10 percent for Microsoft Hyper-V (Windows Server 2008 R2/2012)

  - 5 percent for VMware vSphere (5.*x*)

- Dedicate and reserve CPU and memory resources for the mailbox virtual machines. Do not over-commit these resources.

- Minimize vCPU to pCPU oversubscription to be no larger than a 2:1 ratio; best practice is 1:1.

- Do not over-commit memory.

• Exchange Server virtual machines, including Exchange mailbox virtual machines that are part of a DAG, can be combined with host-based failover clustering (Windows Failover Cluster or vSphere HA cluster).

• In vSphere and Hyper-V, create an anti-affinity rule for each related Exchange virtual machine in the DAG to keep them apart on the hosts within a cluster.

• In vSphere, enable DRS fully automated mode (HA evaluates DRS rules in vSphere 5.5).

• Following the failover (or failback) of a virtual machine from source to target, restart the virtual machine (cold start) on the target.

• Install the latest hypervisor tools on each guest Exchange virtual machine for best performance and specific drivers support: Virtual Machine Integration Services for Hyper-V and VMware Tools for vSphere

• Use both hypervisor and OS performance tools to monitor the virtual machines.

Virtual disk types   In VMware, VMFS supports Raw Device Mapping (RDM), similar to the way Windows Hyper-V supports pass-through disks. RDM, like a pass-through disk, allows a virtual machine to directly access a volume on the physical storage subsystem and can only be used with Fibre Channel or iSCSI.

Base the decision to use specific disk types for Exchange data on technical requirements. Table 13 provides considerations when making disk-type decisions.

Table 13.   Virtual disk type format considerations for Exchange

VMFS/VHD/VHDX	RDM/Pass-through
- Volume can contain many virtual machine disk files, reducing management overhead - Increases storage utilization and provides better flexibility and easier administration and management - Supports existing and future hypervisor storage virtualization features - VMFS fully supports VMware vCenter SRM, vSphere vMotion, vSphere HA, and DRS - Large disk size support - vSphere—VMFS volumes up to 64 TB and virtual disks VMDK files up to 62 TB - Hyper-V—VHDX format up to 64 TB	- Ideal if disks must be dedicated to a single virtual machine - Requires more LUNs, making it easier to reach the limit (Example, max 256 LUNs that might be presented to an physical vSphere host) - Might be required to use array-level backup and replication tools (VSS) integrated with Exchange databases - Facilitates data migration between physical and virtual machines using the LUN swing method - RDMs fully supports vCenter SRM, vSphere vMotion, vSphere HA, and DRS - Supports presenting volumes of up to 64 TB to the guest operating system—Physical-mode RDM only (RDM/P)

Virtual SCSI Adapters

When virtualizing Exchange on VMware, there are options to choose different types of SCSI adapters for a guest virtual machine. The two commonly used virtual SCSI adapters for Windows Server 2008 R2 and Windows Server 2012 are LSI Logic SAS and VMware Paravirtual SCSI (PVSCSI).

- **LSI Logic SAS** is the default adapter when creating new virtual machines with either of these two operating systems. The LSI Logic SAS adapter requires no additional drivers and can satisfy the requirements of most workloads. Because the adapter requires no additional drivers it has been selected as the default vSCSI adapter by many organizations. Queue depth for LSI Logic SAS controller is 32.

- The **VMware Paravirtual SCSI** adapter is a high-performance vSCSI adapter that provides optimal performance for virtualized business-critical applications. The advantage of the PVSCSI adapter is that it delivers added performance while minimizing the use of hypervisor CPU resources. This leads to less hypervisor overhead required to run storage I/O-intensive applications. Queue depth for Paravirtual SCSI controller is 64.

Note: If you are considering using PVSCSI in environments running Exchange on Windows 2008 R2 and ESXi versions prior to 5.0 Update 1, refer to this Knowledge Base article.

Exchange 2013 has greatly reduced the amount of I/O generated to access mailbox data compared to previous versions; however, storage latency is still a factor. In environments supporting thousands of users per Mailbox server, PVSCSI might prove beneficial. Decide whether to use LSI Logic SAS or PVSCSI based on Jetstress testing of the predicted workload using both adapters. Additionally, organizations must consider any management overhead an implementation of PVSCSI might introduce. Because many organizations have standardized on LSI Logic SAS, if the latency and

Microsoft Exchange Server
Best Practices and Design Guidelines for EMC Storage:
EMC VNX Series, EMC Symmetrix VMAX systems

34

throughput difference is negligible with the proposed configuration, the best option might be the one with the least impact to the current environment.

During allocation of storage devices, each device is assigned to a vSCSI adapter in sequential order. A new vSCSI adapter is not created until a vSCSI adapter reaches its 15th device. To provide better parallelism of storage I/O, equally distribute storage devices among the four available vSCSI adapters.

Note. Virtual machines can be deployed with up to four virtual SCSI adapters. Each vSCSI adapter can accommodate up to 15 storage devices for a total of 60 storage devices per virtual machine.

Unsupported configurations

Microsoft *does not* support Exchange with the following:

- NFS shares or SMB implementations other than SMB 3.0

- VMDKs on NFS

- Hypervisor Snaps

- Applications running at the root (applies to Hyper-V only)

- VSS backup of root for pass-through disks—backups must be performed from within the guest virtual machine

- Hyper-V replica

The References section provides more information about Exchange virtualization.

XtremSW Cache with Exchange

Overview

EMC XtremSW Cache is a server flash caching solution that reduces latency and increases throughput to improve application performance by using intelligent software and PCIe Flash technology. XtremSW Cache accelerates reads and protects data by using write-through cache to the networked storage to deliver persistent high availability and disaster recovery. It creates the most efficient and intelligent I/O path from the application to the data store. The result is a networked infrastructure that is dynamically optimized for performance, intelligence, and protection for both physical and virtual environments.

XtremSW Cache benefits

XtremSW Cache provides complete and flexible control over the scope and detail at which it can be enabled. In physical environments, you can enable or disable XtremSW Cache at the source volume level or at the LUN level. In virtual environments, you provision the XtremSW Cache capacity to an individual virtual machine. You then configure the allocated cache capacity inside the virtual machine at the virtual disk level.

The References section on page 45 provides more information.

How does XtremSW Cache accelerate Exchange performance?

When XtremSW Cache is configured for volumes housing Exchange databases, VFCache accelerates block I/O reads that require the highest IOPS and/or the lowest response time. The software uses the PCIe card as a cache of the most frequently referenced data, shrinking storage access time while offloading the I/O processing from the storage array. By residing on the server on the PCIe bus, XtremSW Cache bypasses the overhead of network storage access, thus reducing response times.

XtremSW Cache puts Exchange data into the server I/O stack closer to the application, which dramatically improves performance. Our validation testing of XtremSW Cache with Exchange 2010 shows a significant reduction in user response times and increased throughput. If you have heavy to extreme Exchange workload requirements that are greater than 250 messages per user per day, you should consider implementing an XtremSW Cache solution.

Note: Performance gain and reduction in response times can vary based on each customer's Exchange email usage. We recommend that you use a pilot phase test in your environment to determine the exact benefits of this technology. However, all testing within EMC showed dramatic improvements in performance.
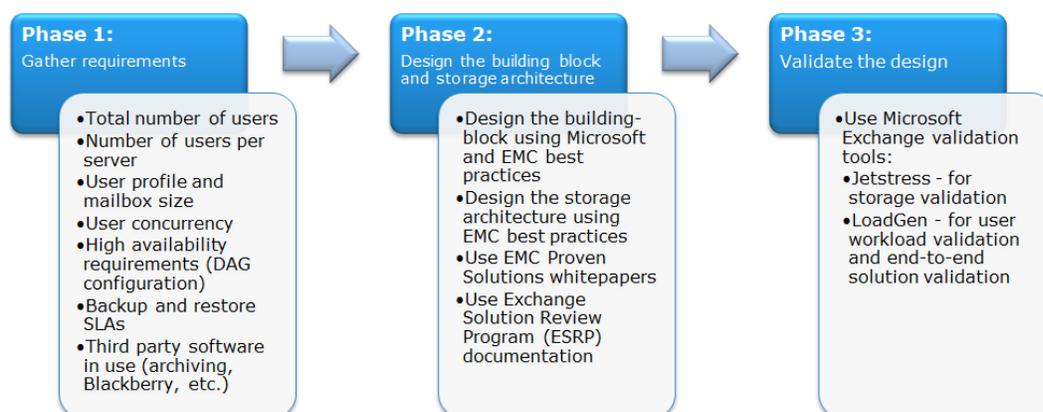
Exchange storage design

Sizing and configuring storage for use with Exchange Server 2010 can be a complicated process driven by many variables and requirements. Properly configured Exchange storage combined with optimally sized server and network infrastructures can guarantee smooth Exchange operation and an excellent user experience. One of the methods that can be used to simplify the sizing and configuration of large amounts of storage on EMC VNX family or Symmetrix VMAX series storage arrays for use with Exchange Server is to define a unit of measure—a *mailbox server building block*.

Mailbox server building block

A mailbox server building block represents the amount of storage (I/O, capacity, and bandwidth), server (CPU, memory), and network resources required to support a specific number of Exchange Server users. The amount of required resources is derived from a specific user profile type, mailbox size, and disk requirements. Using the building block approach simplifies the design and implementation of Exchange Server.

Once you design the initial building block, you can easily reproduce it to support the required number of users in your enterprise. By using this approach, EMC customers can create their own building blocks that are based on their company's specific Exchange environment requirements. This approach is very helpful when future growth is expected because it makes Exchange environment expansion simple and straightforward. EMC best practices involving the building block approach for Exchange Server design have proven to be very successful in many customer implementations.

Designing a building block that is appropriate for a specific customer's environment involves three phases: Collect the relevant requirements, build the block, and validate the design. Figure 2 illustrates the Exchange building block design process.

Figure 2.   **Phases in Exchange building block design**

Appendix A: Exchange manual storage sizing process provides an example of designing an Exchange Server building block on EMC storage by using a three-phase approach. The focus is on the storage aspect of designing an Exchange Server building block—specifically storage I/O, capacity, and bandwidth.

Note: Design guidelines for server resources (such as CPU and memory) and network design guidelines are beyond the scope of this document.

Storage design validation

The final step in the Exchange pre-production deployment phase is to validate the storage to ensure that it is configured properly and can sustain the load it is designed to support.

Jetstress

The Microsoft Jetstress tool is used to validate the Microsoft Exchange storage design. The tool simulates Microsoft Exchange I/O at the database level by interacting directly with the Extensible Storage Engine (ESE) database technology (also known as Jet), on which Microsoft Exchange is built. Jetstress can be configured to test the maximum I/O throughput available to the disk subsystem within the required performance constraints of Microsoft Exchange. Jetstress can accept a simulated profile of specific user count and I/O per second (IOPS) per user to validate that the disk subsystem is capable of maintaining an acceptable performance level by generating the metrics defined in that profile. EMC recommends that you use Jetstress testing to validate storage reliability and performance prior to the deployment of the Microsoft Exchange production environment.

Jetstress is available for download from the following Microsoft sites:

- Jetstress 2010:
  http://www.microsoft.com/en-us/download/details.aspx?id=4167

- Jetstress 2013:
  http://www.microsoft.com/en-us/download/details.aspx?id=36849

The References section on page 45 provides more information.

ESRP Solutions

The Exchange Solution Review Program (ESRP) is a Microsoft program for validating storage vendors' Exchange Server designs. Vendors run multiple Jetstress tests based on various performance, stress, backup-to-disk, and log file replay requirements. Microsoft reviews the results and posts approved solutions on Microsoft.com. You can view Microsoft approved reports for EMC submitted solutions at http://technet.microsoft.com/en-us/exchange/ff182054.aspx.

EMC Proven Solutions and white papers

EMC reference architectures and white papers on the subject of Exchange Server storage design are available at http://www.emc.com/exchange and http://www.microsoft.com/emc.

High availability and data protection for Exchange Server on EMC storage

In Exchange 2010 and Exchange 2013, Microsoft introduced the DAG feature to enhance its native Exchange high availability and data protection capabilities.

Exchange DAG

A DAG, or database availability group, is a set of Microsoft Exchange Server Mailbox servers that can provide automatic database-level recovery from a database, server, or network failure. Mailbox servers in a DAG (16 maximum) monitor each other for failures. When new Mailbox server is added to a DAG, it works with the other servers in the DAG to provide automatic, database-level recovery from database, server, and network failures. You can extend DAGs to multiple sites and provide resilience against datacenter failures.

EMC has a number of data protection products and options that complement DAG and can help further protect your Exchange environment from the loss of a database, server, or an entire site. This section describes various Exchange server high availability and disaster recovery options.

Lagged copy

If a past point-in-time copy of mailbox data is required, Exchange provides the ability to create a lagged copy in a DAG environment. This can be useful in the rare event that a logical corruption replicates across the databases in the DAG, resulting in the need to return to a previous point in time. Lagged copies can also be useful if someone accidentally deletes mailboxes or user data.

One of the disadvantages of lagged copies is that they cannot provide multiple point-in-time recoveries. After using a lagged copy for recovery, you must create a new lagged copy. This process may be resource and time consuming in large environments. For additional information about Exchange database corruption types and lagged copy limitations, go to http://technet.microsoft.com/en-us/library/dd335158.aspx.

EMC snapshots provide the same or better protection levels using far less storage. EMC replication technologies such as VNX Snapshots, TimeFinder Snaps, and RecoverPoint provide instant point-in-time recovery from a snapshot.

EMC storage-based replication compared to native host-based DAG replication

Exchange 2013 and Exchange 2010 include a third-party replication API that enables organizations to use third-party synchronous DAG replication solutions instead of the built-in continuous DAG replication feature. Microsoft fully supports third-party solutions that use this API, as long as the solution provides the necessary functionality to replace all native continuous replication functionality that is disabled as a result of using the API. More information about third-party APIs can be found in the online Exchange technical documentation at http://technet.microsoft.com/en-us/library/dd638137(v=exchg.150).aspx#API.

EMC is the first company to implement third-party API functionality and allow customers to use their existing SAN and storage-based replication with DAG in third-party mode instead of native host-based replication. EMC implements third-party API replication with EMC Storage Integrator (ESI) for Windows suite. This free product allows Exchange Administrators to configure Exchange and to provision, monitor, and

automate storage tasks from Microsoft management consoles for a wide variety of EMC storage arrays.

This solution provides significant cost and storage savings over using native DAG replication. Using storage-based replication frees up mailbox server (CPU & memory) and network resources when native host-based replication is used. This solution eliminates the need to deploy hundreds of database copies from your Exchange environment, but still provides full protection and high availability to users.

EMC Storage Integrator (ESI) for Windows

EMC allows Exchange administrators to be in control of provisioning and optimizing storage for their Exchange environment with EMC Storage Integrator (ESI) for Windows Suite. The ESI for Windows Suite is a set of tools for Microsoft Windows and Microsoft applications administrators. The suite includes features and functionality that allow you to manage your applications on premises or in the private cloud. Specifically, the ESI Exchange Integration feature enables you to integrate Microsoft Exchange with supported EMC storage systems. This integration combines the ESI Microsoft Exchange Adapter, the ESI Exchange High Availability (HA) Extension, the ESI Exchange SCOM Management Packs, and the Exchange cmdlets in the ESI PowerShell Toolkit. In addition, ESI provides a utility for optimizing VNX pool LUNs for Exchange workload. You can connect and manage Microsoft Exchange storage as follows:

- Discover Exchange mailbox databases, mailbox database copies, mailbox servers, and DAGs

- Associate Exchange objects with applicable EMC storage-system objects

- Create and remove Exchange mailbox databases

- View database copy status in the ESI GUI or with the ESI PowerShell Toolkit

- Add and remove mailbox database copies for high availability and site resilience

- Create, provision, and remove storage for mailbox databases

- Use the ESI PowerShell Toolkit to automate storage-provisioning and high availability tasks

- Use the ESI for VNX Pool Optimization utility to improve performance of storage pool-based thin and thick LUNs for Jetstress pre-deployment validation

- Use EMC RecoverPoint to enable Exchange third-party DAGs (in third-party replication mode) for high availability

- Use the ESI Exchange SCOM Management Packs to monitor the ESI Exchange HA Extension in SCOM. You can use SCOM to ensure that the extension is running and to generate alerts for possible remediation failures.

VNX Snapshots

VNX Snapshots is a VNX software feature that creates point-in-time data copies that customers can use for data backups, software development and testing, repurposing, data validation, and local rapid restores. Depending on your application requirements, you can create VNX Snapshots and snapshots of VNX Snapshots. VNX Snapshots support 256 writable snaps per LUN and up to 32,000 per system, depending on the VNX system size. VNX Snapshots do not replace SnapView technology. In fact, both features can coexist. VNX Snapshots supports Pool LUNs,

while VNX SnapView supports RAID Group LUNs. In addition, SnapView clones can have their own VNX Snapshots. Unlike SnapView snapshots and clones, VNX Snapshots do not consume large amounts of pool capacity. As a result, this feature is preferred for modern data management.

VMAX TimeFinder Snaps

VMAX TimeFinder software creates snapshots that are virtual point-in-time copies of a LUN. Snaps are comprised of data that resides on the source LUN and on the save devices. The data on the save devices consists of copies of the original source LUN data that have been modified since the snap was taken. When a snap is taken, the production host is still able to write to the source LUN and modify data. When this occurs, the software stores a copy of the original data in the save devices. This operation is referred to as copy on first write and occurs only once for each chunk of data that is modified on the source LUN. TimeFinder snaps have minimal to no impact on the Exchange servers during backups. When used in combination with EMC Replication Manager, snapshots are application-consistent because they are created within the VSS framework.

EMC data protection offerings

EMC data protection offerings for Exchange (with VNX Snapshots or RecoverPoint) are:

- EMC Replication Manager (EMC RecoverPoint, EMC SnapView, EMC TimeFinder)
- EMC AppSync (VNX Snapshots, EMC RecoverPoint)
- EMC Avamar
- EMC Networker®

Each of these listed solutions has its own strengths and capabilities. The choice depends on the individual customer's requirements for backup and restore and what type of EMC storage array they currently have or are planning to acquire in the future.

EMC hardware-based snapshot and clone products have been integrated with Microsoft VSS technology for many years. TimeFinder, VNX SnapView, and VNX Snapshots enable you to create local point-in-time snapshots or data clones for backup and recovery operations. These products enable simple, non-destructive backup operations with space-saving snapshots or full block-for-block clone copies of your databases and logs. With these products, backups and restores can occur in seconds.

EMC Replication Manager

EMC Replication Manager enables the management of EMC point-in-time replication technologies through a centralized management console for both VNX and VMAX systems. Replication Manager coordinates the entire data replication process—from discovery and configuration to the management of multiple, application-consistent, disk-based replicas.

Replication Manager uses the VSS functionality provided by Microsoft that facilitates the creation of application-integrated snapshot backups. Specifically, Replication Manager provides support for Exchange Server snapshots using Volume Shadow Copy Service (VSS). Replication Manager supports Exchange Server 2010/2013 in standalone or DAG environments.

EMC AppSync

EMC AppSync offers a simple, self-service SLA-driven approach for protecting virtualized Microsoft applications in VNX deployments. After defining service plans, application owners can protect production data and recover data quickly with item-level detail. AppSync also provides an application protection monitoring service that generates alerts when the SLAs are not met.

The *Advanced Protection for Microsoft Exchange2010 on EMC VNX Storage white paper* provides a complete solution using Exchange with EMC AppSync and VNX snapshots. The References section on page 45 provides more information.

Rapid Exchange backup and restore

EMC strongly recommends a robust method of enabling rapid Exchange database backup and restore. EMC AppSync, EMC Replication Manager, EMC Avamar, and EMC Networker offer features for log truncation and the mounting of databases to alternative hosts.

Even if the native Microsoft Exchange DAG feature is used with or without a lagged copy, EMC strongly recommends an alternative, solid, point-in-time Exchange data protection strategy to guard against logical corruption events.

EMC ItemPoint

In EMC Replication Manager or EMC AppSync, there is an option to install EMC ItemPoint. EMC ItemPoint for Microsoft Exchange Server works with your existing Microsoft Exchange Server backup architecture and procedures and allows you to recover individual mailboxes, folders, messages, attachments, calendar items, notes, and tasks directly to your production Microsoft Exchange Server or any Personal Folders (PST) file. This eliminates time-consuming and expensive single mailbox (brick-level) backups. This powerful software also lets you search and create a copy of all archived email that matches a given keyword or criteria.

The References section on page 45 provides more information.

Disaster recovery options for Exchange

EMC offers various disaster recovery (DR) options for Exchange Server. Table 14 presents the three most popular options in addition to the native Exchange DAG feature. Each option has its advantages and disadvantages. Your specific DR requirements determine the best option for you.

Table 14.   EMC disaster recovery offerings for Exchange 2010

Offering	Replication method	Description
Cross-site (stretch) DAG	Exchange continuous replication (DAG)	Built into Exchange 2010/2013 for high availability and disaster recovery
Database portability	EMC RecoverPoint	Only Exchange data is replicated. Users are re-homed to a different server at a secondary site.
Server/site move	EMC RecoverPoint, EMC SRDF	Both the OS and Exchange data are replicated. Failover includes server start, IP change, and DNS update. The entire process is automated using vSphere with Site Recovery Manager.
EMC Storage Integrator	EMC RecoverPoint	• Synchronous or asynchronous solution • Integrates with Exchange 2010/2013 third party replication API • Replaces native DAG replication

Additional backup recommendations

For Exchange backup, EMC recommends the following:

- With medium to high user I/O profiles, to reduce performance degradation, do not take backups directly from the production Exchange server. Instead, mount a point-in-time snapshot or clone on a different server and take the backup from that server.

- Use passive copies for backups whenever possible.

- Schedule backups to take place during off hours, whenever possible.

Conclusion

This document highlights the key decision points in planning a Microsoft Exchange Server deployment with EMC storage systems. Multiple configuration options are available to suit most customer requirements. EMC storage and data management software products are designed to provide the flexibility customers need to manage their Exchange Server environments in a manner that best meets their business requirements.

Best practices for designing Exchange Server storage are constantly evolving. This document presents a snapshot of the current best practices recommended by EMC for deploying Exchange Server with the EMC VNX family of unified storage or EMC Symmetrix VMAX series storage systems. Following these guidelines can assist you in achieving an efficient, high-performance, and highly available Exchange Server environment that meets the customer's requirements.

This paper presents concepts, principles, and formulas to help you:

- Understand the I/O and bandwidth characteristics of Exchange Server

- Apply VNX- and VMAX-specific best practices for Exchange Server

- Utilize an Exchange Server storage building block

- Calculate storage I/O, capacity, and bandwidth requirements

- Validate your overall storage design

- Become familiar with various data protection options for Exchange Server

Additional information

For additional guidance on deploying Microsoft Exchange Server with EMC VNX family of unified storage or EMC Symmetrix VMAX series storage, consult your local EMC Microsoft Exchange expert.

EMC²®

Microsoft Exchange Server   44
Best Practices and Design Guidelines for EMC Storage:
EMC VNX Series, EMC Symmetrix VMAX systems

References

EMC white papers   For additional information, see the white papers listed below.

- *EMC VNX Virtual Provisioning*
- *Implementing EMC Symmetrix Virtual Provisioning With VMware vSphere*
- *EMC FAST VP for Unified Storage Systems*
- *EMC Symmetrix VMAX with Enginuity*
- *Accelerating Microsoft Exchange 2010 Performance with EMC XtremSW Cache*
- *Advanced Protection for Microsoft Exchange2010 on EMC VNX Storage*
- *VNX Snapshots*

Product documentation   For additional information, see the product documentation listed below.

- *EMC AppSync*
- *EMC XtremSW Cache*
- *EMC Storage Integrator (ESI) for Windows Suite*
- *EMC Replication Manager*
- *EMC VNX Family*
- *EMC VNX Series Specifications*
- *EMC Symmetrix VMAX Family*
- *EMC ItemPoint*
- *EMC Storage Integrator for VNX Pool Optimization Release Notes*
  (Available on EMC Online Support mid-November 2013)

Other documentation   For additional information, see the documents listed below.

- *The Essential Guide to Exchange and Private Cloud*, by Paul Robichaux
- *Microsoft Exchange 2010 on VMware Best Practices Guide*
- *Understanding Exchange 2010 Virtualization*
- *Best Practices for Virtualizing Exchange Server 2010 with Windows Server 2008 R2 Hyper-V*
- *Best Practices for Virtualizing and Managing Exchange 2013*
- *Jetstress 2010 Field Guide* and *Jetstress 2013 Field Guide*
- *Exchange 2013 Storage Configuration Options*
- *Exchange 2013 Server Role Requirements Calculator*
- *EMC Exchange 2010-2013 Designer*

EMC²®

Appendix A: Exchange manual storage sizing process

Overview

This section describes a proven building block approach for designing an Exchange environment. This approach has helped many customers simplify the design and implementation of Exchange in both physical and virtual environments. EMC's experience and proven building block methodology ensures a predictable, reliable, scalable, and high-performance design.

The example presented in this section provides instructions for designing an Exchange Server 2013 building block on EMC storage by using the three-phase approach described in Exchange storage design.

The focus is on the storage aspect of designing an Exchange Server building block—specifically storage I/O, capacity, and bandwidth. Design guidelines for server resources (such as CPU and memory) and network design guidelines are beyond the scope of this document.

Requirements gathering

Phase one of the design process involves collecting all relevant customer requirements. The importance of this phase cannot be overstated. Put the requirements in writing and have all parties agree on them. If the requirements change, inform all relevant parties of the changes and assess the potential impact of the changes on the project.

Key Exchange Server requirements

Key requirements for Exchange Server 2010 and 2013 include the following:

- Total number of users (mailboxes) in the Exchange environment
- Number of users per mailbox server
- User I/O profile (number of messages sent/received per day)
- Mailbox size limit
- Read/write ratio
- Average message size
- Outlook client mode (online or cached)
- Log protection buffer
- Deleted items retention policy
- User concurrency requirements
- If replication is needed, the number of DAGs and database copies required
- Backup and restore requirements (RTO/RPO)
- Third-party software that affects space or I/O (for example, Blackberry, snapshots)
- Storage infrastructure
- Physical or virtual deployment

Table 15 presents an example of sizing the customer requirements for a large Exchange Server 2013 deployment.

Table 15. **Example of customer requirements for Exchange deployment**

Item	Value
Total number of active users (mailboxes) in Exchange environment	20,000
Site resiliency requirements	Single site
Storage Infrastructure	SAN
Type of deployment (Physical or Virtual)	Virtual (VMware vSphere)
HA requirements	One DAG with two database copies
Mailbox size limit	2 GB max quota
User profile	200 messages per user per day (0.134 IOPS)
Target average message size	75 KB
Outlook mode	Cached mode, 100 percent MAPI
Number of mailbox servers	8
Number of mailboxes per server	5,000 (2,500 active/2,500 passive)
Number of databases per server	10
Number of users per database	500
Deleted items retention (DIR) period	14 days
Log protection buffer (to protect against log truncation failure)	3 days
BDM configuration	Enabled 24 x7
Database read/write ratio	3:2 (60/40 percent) in a DAG configuration
User concurrency requirements	100 percent
Third-party software that affects space or I/O (for example, Blackberry, snapshots)	Storage snapshots for data protection
Disk Type	3 TB NL-SAS (7,200 rpm)
Storage platform	VNX

Exchange Mailbox server sizing tools

After collecting user requirements, you can begin the storage and server sizing process.

Microsoft and EMC provide tools to help you properly size your Exchange Mailbox server. The Microsoft *Exchange 2013 Mailbox Server Role Requirements Calculator* tool provides CPU and memory guidance in addition to storage recommendations. EMC enhances Microsoft's calculator and adds special EMC extensions that are specifically designed and tailored for EMC VNX and VMAX storage. EMC's enhanced

calculator provides more in-depth details and accurate recommendations for deploying Exchange 2010 and Exchange 2013 on EMC storage. Specifically, the *EMC Exchange 2010-2013 Designer* includes features such as thin provisioning and FAST VP for sizing Exchange storage.

For access to the *Microsoft Exchange 2010 Mailbox server Role Requirements Calculator*, visit the Microsoft TechNet at
http://gallery.technet.microsoft.com/Exchange-2010-Mailbox-Server-Role-/

For access to the *Microsoft Exchange 2013 Mailbox Server Role Requirements Calculator*, visit the Microsoft TechNet library at
http://gallery.technet.microsoft.com/Exchange-2013-Server-Role-f8a61780

For access to the *EMC Exchange 2010-2013 Designer*, visit "Everything Microsoft at EMC" section on EMC Community Network portal at
https://community.emc.com/docs/DOC-13037.

Consult EMC for any additional guidelines during the design and deployment phases.

While the Microsoft and EMC calculators are both viable options to start the initial design sizing's modeling exercise, many experienced Exchange and Storage administrators prefer to do the calculation themselves. The following section presents an example of how to calculate disk requirements based on user IOPS and capacity requirements for Exchange server 2013.

The process is based on information provided in the Exchange Team Blog technical article Sizing Exchange 2013 Deployments.

Defining a building block

The Exchange storage design phase involves defining a building block that satisfies the collected requirements. The building block includes storage, server, and network resources for a single mailbox server. This exercise focuses on defining the storage resources. To size the storage requirements, calculate the IOPS requirement first, and then calculate the capacity requirement. Determine the total number of disks required by consolidating the results from the IOPS and capacity requirement calculations.

In our example, we are defining a building block as a mailbox server that will support 5,000 users, where 2,500 users will be active during normal runtime and the other 2,500 users will be passive until a switchover from another mailbox server occurs. Each building block will support two database copies.

Exchange storage IOPS calculations

The first step in determining the storage resources required for a building block is to calculate disk requirements based on user I/Os. In our example, each mailbox server needs to support 5,000 users (2,500 active and 2,500 passive) with a 200-messages-per-user-per-day profile, which corresponds to 0.134 IOPS per user. Then, scale up to eight additional building blocks to accommodate configuration with 20,000 active users in a DAG with two database copies.

To size the Exchange storage IOPS you need to calculate database and log volume IOPS requirements. You need to know how many Exchange IOPS each disk type can handle. Based on example requirements, we used NL-SAS disks on VNX storage.

Using values from Table 11 on page 18, the NL-SAS disk can deliver 65 random and 180 sequential IOPS with an Exchange 2010/2013 workload.

Database IOPS Requirements

To calculate the database volume IOPS requirements, complete the following steps:

1. Calculate the total database IOPS (total transactional IOPS) required to support all users (active and passive) on a single mailbox server (one building block) by using the following calculations. This will be the target IOPS value you must achieve during Jetstress validation:

    ```
    Total transactional IOPS = IOPS per mailbox * mailboxes
    per server + (Microsoft recommended overhead %)
    ```

    Note: Microsoft typically recommends a 20 percent I/O overhead factor to account for special customer requirements and future growth.

    Example:

    ```
    Total transactional IOPS = 5,000 users * 0.134 IOPS per
    user + 20% Microsoft recommended overhead = 670 + 134 =
    804 IOPS
    ```

2. Calculate front-end IOPS by adding any additional overhead values required by EMC:

    ```
    Total front-end IOPS = (Total transactional IOPS) + (EMC
    required overhead %)
    ```

    Note: Do not confuse the additional EMC required 20 percent IOPS overhead with the Microsoft recommended 20 percent overhead, which the individual customer may or may not choose to add.

    Example:

    Total Front-end IOPS `= 804 + 20% EMC required overhead` `=` **965 IOPS** `(rounded-up from 964.8)`

Database Disks requirements based on IOPS

Next, calculate the number of disks required to provide the desired user performance for the Exchange database, based on the IOPS requirements. For that, we need to determine back-end IOPS. The back-end IOPS are the IOPS seen by the storage array and include front-end exchange host IOPs plus RAID write penalty.

Use the following formula to calculate database disk requirements based on IOPS:

```
Disks required for Exchange database IOPS = (Total backend
database Read IOPS) + (Total backend database Write IOPS)/
Exchange random IOPS per disk
```

```
Where:
Total back-end database read IOPS = (Total Front-end IOPS) * (% of
Reads IOPS)
```

```
Where:
Total back-end database write IOPS = RAID Write Penalty *(Total
Front-end IOPS * (% of Write IOPS)
```

Note: For Exchange IOPS per disk, use values from Table 11. For read-to-write ratios use values from Table 7.

Table 16 provides a summary and calculations for disk requirements based on IOPS with each RAID type.

Note: Since each RAID type has a different RAID overhead value, each RAID type yields significantly different results for the number of disks required. It is important that you choose the RAID type that best fits your requirements for I/O and capacity.

Table 16.   Calculating RAID disk requirements

RAID Option	RAID Penalty	Disks required
RAID 1/0 (4+4)	2	(965 x 0.60) + 2(965 x 0.40) = 579 + 772 = 1351 / 65 = 21 (round-up to **24 disks**)
RAID 5 (4+1)	4	(965 x 0.60) + 4(965 x 0.40) = 579 + 1544 = 2123 / 65 = 33 (round-up to **35 disks**)
RAID 6 (6+2)	6	(965 x 0.60) + 6(965 x 0.40) = 579 + 2316 = 2895 / 65 = 45 disks (round-up to **48 disks**)

Based on the above calculations, RAID 1/0 requires the lowest number of disks and is the best choice from an I/O perspective. Nevertheless, this might not be the best choice from capacity standpoint. The final number will be chosen after capacity requirement calculations are completed.

Transactional logs IOPS requirements

Exchange Log volume I/O is sequential in nature with minimal overhead and is not a significant factor for Exchange storage sizing. Nevertheless, it is a best practice to calculate the log IOPS requirements to ensure the overall design can provide good performance.

For Exchange 2010 and 2013 the log-to-database write is 50 percent for databases that participate in mailbox resiliency (DAG) and 40percent for stand-alone deployments. In addition, on mailbox servers that are in the DAG, there is an overhead associated with using continuous replication. This overhead is an additional 10 percent in log reads for each active database copy that is hosted on the Mailbox server. In addition, Microsoft recommends applying a 20 percent I/O overhead factor to ensure adequate room for busier-than-normal user activity periods.

Calculate the number of disks required to provide the desired user performance for Exchange log files, based on the IOPS requirements, by using this formula:

```
Disks required for Exchange log IOPS = (Total backend database
Write IOPS * 50%) + (Total backend database Write IOPS * 10%)/
Exchange sequential IOPS per disk
```

Note: Because we already calculated database back-end IOPS that includes 20 percent Microsoft recommended I/O overhead, we do not need to add it again for transactional logs.

Using our example, to support an Exchange log IOPS requirement for a building block of 5,000 users at 0.134 IOPS per mailbox server using RAID 1/0 and 7,200 rpm NL-SAS drives, we need 4 disks:

Example:

```
(772 back-end write IOPS * 50%) + (772 *10%))/ 180 sequential
Exchange IOPS per disk = (386 + 77.2)/180 = 2.57(round-up to 4
disks)
```

Exchange storage capacity calculations

In the previous steps, we have determined how many disks are necessary to support database and log IOPS requirements for a 5,000-user building block. Now we can calculate how many disks are necessary to support the capacity requirements for the same building block. After we complete the calculations, we can determine whether the number of disks based on IOPS requirements is sufficient compared to the capacity requirement.

Considerations when sizing using thin provisioning

As mentioned in previous sections, you can deploy Exchange on thick or thin LUNs. If you are designing an Exchange solution to use thin provisioning, you must perform two sets of calculations for determining the disk capacity requirements:

- The first set of calculations determines the initial capacity requirements. This is necessary to identify the storage requirements needed to support the initial mailbox capacity. You will base the storage capacity purchase on these calculations.

- The second set of calculations determines thin provisioned capacity requirements. This is necessary to properly configure the size of the database and log LUNs to be presented to the host. This is also necessary for provisioning the required storage for a fully provisioned mailbox.

When designing Exchange using thin provisioning on EMC storage, it is a best practice to separate the Exchange transactional logs from the database using separate storage pools. Since log volumes do not have the same growth pattern as the database volumes, it makes sense to separate them. This also provides the flexibility to put log volumes on different disk types or different RAID levels than the database volumes. This also provides the flexibility to provision Exchange logs on thick or thin LUNs.

To calculate the Exchange storage capacity requirement we need to determine database and log volumes requirements.

To calculate database volume requirements, complete the following steps:

1. Calculate the user mailbox size on disk using this formula:

    Mailbox size on disk = Maximum mailbox size + White space + Dumpster

2. Calculate database size on disk using this formula:

```
Database size on disk = number of mailboxes per database *
mailbox size on disk
```

3. Calculate database LUN size using this formula:

```
Database LUN size = Number of mailboxes * Mailbox size on
disk * (1 + Index space + additional Index space for
maintenance) / (1 + LUN free space)
```

The following sections provide details for all the steps.

Mailbox size on disk calculations

To calculate the mailbox size on disk, use the following formula:

Mailbox size on disk = (Mailbox quota) + (database white space) + (Recoverable items folder size)

- Where **mailbox storage quota** = maximum amount of data that the user can store in their mailbox on the server. In our example, it is 2 GB.

- Where **the database white space** = amount of space in the mailbox database file that has been allocated on disk but does not contain any in-use database pages. The Microsoft recommendation for white space planning is to use the amount of messaging content equal to one day. Use the following formula to estimate database white space:

```
Estimated Database Whitespace per Mailbox = per-user daily
message profile * average message size
```

- Where **recoverable items folder size** (also called "dumpster") = the deleted emails from a user mailbox that are moved temporarily to the recoverable items folder for the duration of the deleted item retention period. Use the following formula to calculate the dumpster space:

```
Dumpster = (per-user daily message profile * average
message size * deleted item retention window) + (mailbox
quota size * 0.012) + (mailbox quota size * 0.03)
```

Example:

White space = 200 messages /day * 75KB = 14.65MB

Dumpster = (200 messages/day * 75KB * 14 days) + (2GB * 0.012) + (2GB x 0.03) = 205.1 + 24.58 + 61.44= 291.12MB

Now that we have all the variables, we can calculate mailbox size on disk using variables from our example:

User mailbox size on disk = 2GB mailbox quota + 14.65MB database whitespace + 291.12MB Dumpster = 2,354 MB (2.3GB)

Database size on disk calculations

Once the mailbox size on disk, whitespace, and dumpster values are determined, we can calculate the database size on disk and database LUN size.

Database size on disk = 500 users per database * 2,354 MB mailbox on disk = 1,177 GB (1.15 TB)

Database LUN size calculations

To determine database LUN size, you must consider the following variables:

- Database size
- Content index size
- Free LUN space requirements

To determine the total database LUN size, use the following formula:

```
Database LUN size = Database size on disk * (1 + Index space +
additional Index space for maintenance)/1 + LUN free space)
```

Content Indexing space considerations

In Exchange 2010 the space for content indexing was estimated at about 10 percent of the database size. In Exchange 2013 this value is now increased to 20 percent. In addition to the already estimated 20 percent per database content indexing space, you must add an additional 20 percent to allow content indexing maintenance tasks (such as the master merge process) to complete.

In our example, considering 20 percent for the Index, 20 percent for the additional Index space for maintenance and 20 percent for LUN-free space protection, the database LUN size will be 1.6TB:

Database LUN size = 1,177 GB * (1 + 0.2 + 0.2) / (1 - 0.2) = 2,060 (2 TB)

Log LUN size calculations

When configuring disks for Exchange, database LUNs are generally given the most attention because it is often thought that database LUNs alone pose the highest risk for performance bottlenecks. It is essential to keep in mind, however, that database transactions are gated by the completion of their associated log writes. Therefore, log LUNs must be given as much consideration as database LUNs during Exchange disk configuration.

You can estimate the storage capacity requirements for a log LUN by considering the following factors:

- **Number of log files generated per mailbox per day**—This depends on the number of emails a user sends and receives per day on an average basis and the average mailbox size. Table 10 provides details.

- **Number of days' worth of logs to maintain**—This depends on how often the log is truncated. Exchange backups, snaps, or clones, which are typically run nightly, truncate log files. Differential backups, snapshots, and clones do not truncate the log files. If the backup/snap/clone schedule includes only weekly

full and daily differential backups/snaps/clones, the log LUN space needs to be larger than an entire week of log file space to accommodate both backup and replay data during a restore operation.

- **Storage required for mailbox moves**—Many large companies move a percentage of their user mailboxes nightly or weekly to different databases, servers, or sites. When a mailbox is moved from one database to another, transaction logs are generated on the target log LUN and are roughly the same size as those generated on the source.

- **Log growth overhead factor**—For most deployments, EMC recommends adding an overhead space of 20 percent to the log size when you create the log LUNs (after all of the other factors have been considered) to ensure that the required space is available in the event of unexpected log generation.

- **High availability factors**—The implementation of high availability features, such as DAG, increases storage requirements.

To calculate Log LUN size, use the following formula:

```
Log LUN size = (Log size)*(Number of mailboxes per
database)*(Backup/truncation failure tolerance days)+ (Space to
support mailbox moves)/(1 + LUN free space)
```

To ensure that the mailbox server does not sustain any outages because of space allocation issues, size the transaction log LUNs to accommodate all of the logs that will be generated during the backup set. If the architecture uses the mailbox resiliency and single item recovery features as the backup architecture, allocate enough log capacity to accommodate any backup or truncation failure (for example, a failed database copy prevents log truncation from occurring). In our example, the backup/truncation failure tolerance window is three days.

Using requirements from our example with a 200-messages-per-user-per-day profile, 40 logs per user per day, a 3-day log protection buffer, 1 percent of mailbox moves per week, and 20 percent free LUN space, the log LUN space for 500 users per database is calculated as follows:

Log Capacity to Support 3 Days of Truncation Failure = (500 mailboxes/database x 40 logs/day x 1MB log size) x 3 days = 58.59GB

Log Capacity to Support 1% mailbox moves per week = 500 mailboxes/database x 0.01 x 2.3GB mailbox size = 11.5GB

Log LUN size = 58.59GB + 11.5 GB /(1 - 0.2) = 87.61 GB (round-up to 88 GB)

Total capacity required per building block

After calculating database and log capacity requirements, we can calculate the total capacity requirements for a building block of 5000 users per server. To calculate the total building block capacity requirements, use the following formula:

```
Total LUN size capacity required per server = (database LUN
size per server) + (log LUN size per server) * (number of
databases per server)
```

Table 17 presents the summary of total LUN capacity requirements for the building block of 5,000 users per server.

Table 17.   Total space requirements per 5,000-user building block:

LUN Capacity Type	LUN Capacity Required per server
Database LUN capacity	2,060 GB per LUN * 10 LUNs per server = 20,600 GB
Log LUN capacity	88 GB per LUN * 10 LUNs per server = 880 GB
Total LUN capacity per server	**20,600 + 880 = 21,480 GB**

Total number of disks required

To determine the number of disks required per building block, use the following formulas.

For database disks:

```
Disks required for Exchange database capacity = Total database
LUN size / Physical Disk Capacity * RAID Multiplication Factor
```

For log disks:

```
Disks required for Exchange log capacity = Total log LUN size)
/ Physical Disk Capacity * RAID Multiplication Factor
```

When calculating the number of disks required, use the Raid Multiplication Factor from the following table:

RAID multiplication factors:

RAID Type	RAID Multiplication Factor
RAID 1/0	2
RAID 5 (3+1)	1.33
RAID 5 (4+1)	1.25
RAID 5 (5+1)	1.2
RAID 5 (6+1)	1.17
RAID 5 (7+1)	1.14
RAID 5 (8+1)	1.125
RAID 6 (6+2)	1.33

Calculate database disk requirements for each RAID type to evaluate which one provides most optimal results. In our example, each 3 TB NL-SAS disk provides 2,794.5 GB of raw capacity.

Table 18.   Exchange database disks requirements options for capacity

RAID Option	Disks required
RAID 1/0 (4+4)	20,600/2794.5 * 2 = 7.37 * 2 = 14.74 (round-up to 16 disks)
RAID 5 (4+1)	20,600/2794.5 * 1.25 = 7.37 * 1.25 = 9.2 (round-up to 10 disks)
RAID 6 (6+2)	20,600/2794.5 * 1.33 = 7.37 * 1.33 = 9.8 (round-up to 16 disks)

Calculate log disks requirements using RAID 1/0 as a best practice with 3 TB NL-SAS disks:

Table 19.   Exchange logs disks requirements for capacity

RAID Option	Disks required
RAID 1/0 (1+1)	880 / 2,794.5 * 2 = 0.63 (round up to 2 disks)

Final storage calculation results

At this stage, compare the results from the performance (IOPS) and the capacity calculations and select the best option that meets both requirements. You should select the larger number to meet both requirements within the same RAID type and select the smaller number when comparing different RAID types. In our example, the best option would be RAID 1/0 because it needs only 24 by 3 TB 7,200 rpm NL-SAS disks for Exchange databases and four by 3 TB 7,200 rpm NL-SAS disks for Exchange log files. Table 20 shows a summary of the results:

Table 20.   Final Exchange storage requirements per building block

Volume Type	RAID Option	Disks Required for Performance (IOPS)	Disks Required for Capacity	Best Option
Exchange Databases	RAID 1/0 (4+4)	24 disks	16 disks	24 disks
	RAID 5 (4+1)	35 disks	10 disks	
	RAID 6 (6+2)	48 disks	16 disks	
Exchange Logs	RAID 1/0 (1+1)	4 disks	2 disks	4 disks
Total disks per building block				**28 disks**

Summary of disk requirements

The best configuration, based on the example requirements and calculated values, is 24 by 3 TB NL-SAS disks in RAID 1/0. This configuration meets both IOPS and capacity requirements for a building block of 5,000 users with a 0.134 IOPS user profile and a 2 GB mailbox size limit.

Once the number of disks per building block is established, multiply that number by the required number of building blocks (eight in this example) to obtain the total

number of disk required for the entire 20,000-user Exchange environment. Following the example, we need 224 3-TB NL-SAS disks to satisfy capacity and performance for the entire environment with two DAG copies:

```
Total number of disks required for entire 20,000 users solution in
a DAG with two copies = 28 disk per building block * 8 building
blocks = 224 disks total
```

Storage bandwidth requirements

As mentioned previously, the bandwidth requirements for BDM are greatly reduced in Exchange 2013 compared to previous Exchange 2010 version. Nevertheless, if more than 100 Exchange databases are to be deployed on a single storage array, EMC recommends that you perform bandwidth validation to ensure that the array buses do not become a bottleneck.

The bandwidth validation process involves the following steps:

1. Determine how many databases the customer requires.

2. Ensure that the database LUNs are evenly distributed among the back-end buses and storage processors.

3. Determine the bandwidth requirements per database.

4. Determine the required bandwidth requirements per array bus.

5. Determine whether each bus can accommodate the peak Exchange database bandwidth.

Bandwidth requirements per database

Bandwidth per database is the throughput that *each* database is to generate. To determine the throughput MB/s per database, use the following formula:

```
Bandwidth per database (MB/s) = Total transactional IOPS per
database * 32 KB + Estimated BDM Throughput per database (MB/s)
```

```
Where: 32 KB is an Exchange page size
Where: Estimated BDM throughput per database is 7.5 MB/s for
Exchange 2010 and 2.25 MB/s for Exchange 2013
```

Required throughput calculations

Required throughput (MB/s) per bus is the throughput capability required to support *all* of the databases assigned to a *single* bus. Use the following formula to calculate the required throughput per bus.

```
Required throughput MB/s per bus = (throughput MB/s per
database) * (total number of active and passive databases per
bus)
```

Next, compare the total database throughput requirement (the throughput capability required to support all of the databases assigned to a single bus) with the maximum throughput specification for the bus.

Note: Contact your local EMC Speed Certified Guru to obtain the bus specifications for the specific array you are using, because these specifications change over time—or use *Disksizer* for VNX.

Example:

Assumption: 500 users at 0.134 IOP users per database and 200 databases per bus.

```
Total transactional IOPS per database = (500 * 0.134 * 32 KB
= 2.1 MB/s

Throughput per database = 2.1 MB/s + 2.25 MB/s = 4.35 MB/s

Required throughput per bus = 4.35 MB/s * 200 databases per bus
= 870 MB/s
```

Therefore, if the array supports a maximum throughput of 3,200 MB/s per bus, 200 databases can be supported from a throughput perspective.

Finalizing storage design

In our example, Exchange is configured in a DAG with two database copies. For best performance, protection and storage utilization we configured dedicated storage pools for each mailbox server. Each storage pool holds two copies from different mailbox servers. We also separated the Exchange log files into different storage pools. For better storage utilization, we created one storage pool for logs per four mailbox server building blocks.

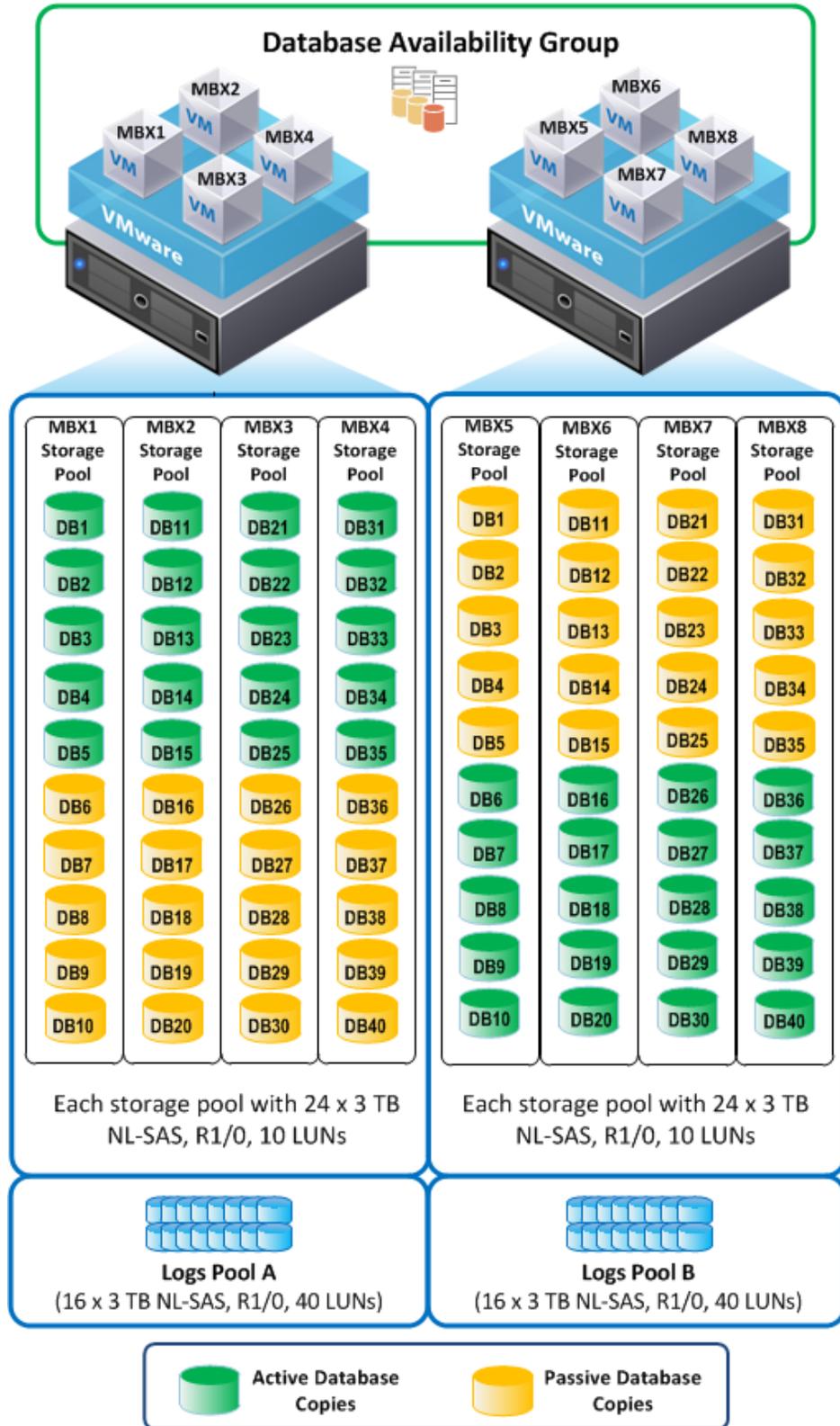Figure 3 provides an example of the Exchange solution storage design on VNX array.

Figure 3.   Exchange storage design example