White Paper

MICROSOFT EXCHANGE 2010 STORAGE BEST PRACTICES AND DESIGN GUIDELINES FOR EMC STORAGE

EMC Solutions Group

Abstract

Microsoft Exchange has rapidly become the choice of messaging for many businesses, and the trend shows no signs of abating. This paper identifies best practices and key decision points for planning and deploying Microsoft Exchange Server 2010 with the EMC® VNX® family of unified storage or EMC Symmetrix VMAX™ series storage.

December 2012

Microsoft Partner
Gold Messaging

EMC²

Part Number h8888.5

Table of contents

Executive summary

In the planning and design phases of a Microsoft Exchange Server 2010 implementation, it is important to understand how the application interacts with the storage platform. It is also critical to know which practices to follow to avoid problems and achieve the best performance.

From a storage design perspective, the application architecture and user profile characteristics of Microsoft Exchange Server 2010 are significantly different from earlier versions of Exchange (2003, 2007).

- One important difference is the introduction of Database Availability Groups (DAGs), which enables multiple copies of a database to reside on different servers.
- Another important difference is the way in which Exchange performs online database maintenance, which can affect performance.
- A reduction in user IOPS, made possible by a new database storage schema, which increases the database I/O page size and the mailbox server cache effectiveness.

In light of these differences, EMC has updated its recommended best practices for Exchange storage design to reflect the features of Exchange Server 2010.

Purpose of this paper

This paper presents the set of current EMC-recommended best practices for storage design in support of Microsoft Exchange Server 2010. Guidelines are presented 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 2010 or considering an upgrade from an earlier version of Exchange. It is assumed that the audience is familiar with Microsoft Exchange Server, EMC VNX family or EMC Symmetrix VMAX series storage, and VMware or Microsoft Hyper-V virtual environments.

Scope

This document presents storage design best practices recommended by EMC for hosting Microsoft Exchange Server 2010 on EMC VNX family storage or EMC Symmetrix VMAX series storage in both physical and 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
Exchange database availability group (DAG)	The base component of the high availability and site resilience framework built into Microsoft Exchange Server 2010. DAG is a set of up to 16 Microsoft Exchange Server 2010 mailbox servers that provide automatic database-level recovery from a database, server, or network failure.
Background database maintenance (BDM)	The process of Exchange Server 2010 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 DAG copies. The process cannot be scheduled, throttled, or disabled on passive DAG copies.
Redundant array of independent disks (RAID)	RAID is 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.
Storage pool	Storage pools are 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.
Thin LUN	A type of pool LUN where physical space allocated can be less than the user capacity seen by the host server.
Thick LUN	A type of pool LUN in which the physical space allocated is equal to the user capacity seen by the host server
FAST Cache	FAST Cache, a VNX storage feature, is software that enables customers to add various Flash drive capacities in order to extend existing cache capacity for better system-wide performance. FAST Cache is now offered with increased capacity configurations using the 100 GB Flash drive or the 200 GB Flash drive.
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.
VFCache	EMC VFCache is a server Flash caching solution anchored in storage array protection that reduces latency and increases throughput to dramatically improve application performance. Intelligent caching software with PCIe Flash technology extends storage tiering to the server.

Featured EMC storage

EMC VNX family of unified storage

The EMC VNX™ family delivers industry-leading innovation and enterprise capabilities for file, block, and object storage in a scalable, easy-to-use solution. 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.

All of this is available in a choice of systems ranging from affordable entry-level solutions to high-performance, petabyte-capacity configurations servicing the most demanding application requirements. The VNX family includes the VNXe™ series, purpose-built for the IT manager in entry-level environments, and the VNX series, designed to meet the high-performance, high-scalability requirements of midsize and large enterprises. For additional EMC VNX system details, visit http://www.emc.com/storage/vnx/vnx-series.htm.

Depending on the platform, two software packs have all the suites supported. For VNXe series, the packs are Total Protection and Total Value. For VNX series, the packs are Total Protection and Total Efficiency packs. The software packs bundle EMC's advanced data management suites:

- **FAST Suite** automatically maximizes capacity and performance efficiency. FAST enables a dynamic FLASH 1st data management strategy where just a small number of Flash drives are used to deliver the highest performance for high activity data, and low activity data is constantly moved to the most cost-effective drive type.
- **Security and Compliance Suite** helps ensure that data is protected from unwanted changes, deletions, and malicious activity. Data is encrypted where it is created for protection anywhere outside the server. File-Level Retention is used to meet compliance requirements. Integration with third-party anti-virus checking, quota management, and auditing applications provides added data protection, security, and peace of mind.

- **Local Protection Suite** combines snapshots and clones with point-in-time recovery with DVR-like rollback capabilities for business continuity on block-based storage, allowing recovery of production applications with minimal data exposure. Application owners can tune recovery point objectives based on criticality of data and perform faster recovery through self-service capabilities. Copies of production data can be used for development, testing, decision support tools, reporting, and backup acceleration.
- **Remote Protection Suite** delivers unified block and file replication, providing disaster recovery for both NAS and SAN environments. It delivers disaster recovery protection for any host and application without compromise—with immediate DVR-like rollback to a point in time. Capabilities include compression and deduplication for WAN bandwidth reduction, application-specific recovery point objectives, and replication options for one-to-many configurations.
- **Application Protection Suite** automates application-consistent copies and enables you to recover to defined service levels. User roles enable self-service copy management, while improving visibility for all application recovery points. Alerts are generated automatically, providing fast resolution to recovery gaps. Integrated reporting can prove compliance with protection policies. Applications supported include Oracle; Microsoft Exchange, SQL Server, and SharePoint; VMware; and Hyper-V.

VNX Snapshots

VNX Snapshots is a new VNX software feature introduced in EMC's VNX operating environment for Block Release 5.32 (OE 5.32). This feature 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 needs, you can create VNX Snapshots and snapshots of VNX Snapshots. VNX Snapshots do not replace the existing 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.

VNX Snapshot technology writes new data to a new area within a storage pool, without the need to read and write the old data block. This improves the overall performance compared to SnapView Snapshots technology.

VNX Snapshots uses redirect on write (ROW) technology. With ROW, the source LUN of a snapshot created by VNX Snapshots does not need to copy source LUN data to the reserved LUN pool to preserve the snapshot. Instead, new incoming writes to the source LUN are written to a location in the storage pool that is different from the location to which the original data is written, and index updates track the data changes.

VNX OE for block 5.32 also introduces the concept of consistency groups. You can combine several pool LUNs into a consistency group and snap them concurrently. When AppSync is used to protect Exchange, there is no need to create consistency groups as AppSync simplifies this process automatically when running a service plan.

New VNX Snapshots now support 256 writable snaps per LUN and up to 32,000 per system (depending on the VNX system size). Snapshots of snapshots make VNX snapshots ideal for testing, development, and disk backups. When used with AppSync, these snapshots are application consistent and can be used to quickly and efficiently provision copies of production data for application development and testing.

You can create a snapshot from:

- Primary LUNs
- Consistency group(s)
- Snapshot mount points (if attached to a snapshot)
- Snapshots (through the Copy Snapshot operation)

Some of the most important uses of VNX Snapshots include:

- Backing up point-in-time copies of your system to a backup server
- Verifying data integrity before restoring data by mounting snapshots
- Restoring data from a local snapshot
- Using a copy of the LUN for testing and development
- Deploying multiple copies of a single source for use in environments such as virtual environments and workstation deployments

More information about VNX snapshot is provided in the *VNX Snapshots* white paper at the www.emc.com website.

EMC Symmetrix VMAX series

The EMC Symmetrix VMAX series is high-end storage for the data center. The system scales to a massive 2 PB and consolidates more workloads with a much smaller footprint than other arrays. EMC Symmetrix® Virtual Matrix Architecture seamlessly scales performance, capacity, and connectivity on demand to meet all application requirements. The system supports Flash Drives, Fibre Channel, and SATA drives, plus optimized automatic tiering with FAST VP. The system also supports virtualized and physical servers, including open systems, mainframe, and IBM hosts.

	VMAX 10K	VMAX 20K	VMAX 40K
Maximum Drives	1,080	2,400	3,200
Usable Capacity	1.5 PB	2 PB	4 PB (3.5" drives) and 2.8 PB (2.5" drives)
Connectivity	FC, 1GbE, 10GbE, 10 Gb/s FCoE, iSCSI	FC, FICON, 1GbE, 10GbE, 10 Gb/s FCoE, iSCSI	FC, FICON, 1GbE, 10GbE, 10 Gb/s FCoE, iSCSI
Front-end Ports	64	128	128
Local Replication	TimeFinder, RecoverPoint	TimeFinder, RecoverPoint	TimeFinder, RecoverPoint
Remote Replication	SRDF, RecoverPoint	SRDF, RecoverPoint	SRDF, RecoverPoint

For additional details about VMAX systems, visit http://www.emc.com/storage/symmetrix-vmax/symmetrix-vmax.htm

Table 3 describes the features and benefits provided in the EMC Symmetrix VMAX series systems.

Table 2. EMC Symmetrix VMAX series features

Feature	Description
Symmetrix FAST and FAST VP	Automate storage tiering to lower costs and deliver higher service levels.
Linear scale-out of storage resources	Consolidate multiple arrays into a single Symmetrix VMAX system.
Up to 2 PB usable capacity	Seamlessly scale from 48 to 2,400 drives.
1 to 8 VMAX engine scaling	Consolidate more workloads in a smaller footprint with up to eight highly available Symmetrix VMAX engines.
Virtual logical unit number (LUN) technology	Transparently move data to the right tiers and RAID types at the right time.
Virtual provisioning	Efficiently allocate, grow, and reclaim storage with ease.
Extended distance protection	Replicate data over extended distances, and achieve zero data loss protection.
Information-centric security	Get advanced RSA security technology—built in, not bolted on—to keep your data safe, reduce risk, and improve compliance.

I/O and bandwidth characteristics of Exchange Server 2010

Microsoft has made significant changes to the Exchange Server 2010 storage schema to accommodate new high availability scenarios and larger mailboxes with fewer I/Os per second (IOPS). It is very important to understand the I/O and bandwidth characteristics of Exchange Server 2010 in order to design an efficient storage configuration. To highlight the changes in user IOPS from previous versions of Exchange, the user profile characteristics of Exchange 2007, Exchange Server 2010 standalone, and Exchange Server 2010 DAG are presented for comparison in Table 3. These estimated measurements include all database volume I/O, such as database, content indexing, and NTFS metadata, but do not include log volume I/O. The estimates are based on an average message size of 75 KB.

Table 3. User profile comparisons across Exchange versions based on an average message size of 75 KB

User profile	Messages sent/received per day	Exchange 2007 user IOPS	Exchange Server 2010 user IOPS for unprotected database copy (standalone)	Exchange Server 2010 user IOPS for protected database copy (mailbox resiliency)
Light	25	0.11	0.040	0.030
Average	50	0.18	0.060	0.050
Heavy	100	0.32	0.120	0.100
Very Heavy	150	0.48	0.180	0.150
Extra Heavy	200	0.64	0.240	0.200
Extreme	500	1.60	0.600	0.500

For more information about the Exchange 2010 memory guidelines, see Understanding Memory Configurations and Exchange Performance on Microsoft TechNet.

Mailbox database I/O read/write ratios

With Exchange 2010, 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/O. If you follow the recommended memory guidelines, you can expect to see the I/O ratios for active database copies displayed in Table 4. Note, that these ratios might be different if you use server cache technologies for your Exchange deployment. For example, EMC FAST Cache or EMC VFCache. These technologies cache more reads, thus change the overall application read to write ratio.

For more information about the memory guidelines, see Understanding Memory Configurations and Exchange Performance on Microsoft TechNet. This measurement includes all database volume I/O (database, content indexing, and NTFS metadata); it does not include log volume I/O.

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 servers, follow the instructions in the section entitled EMC storage design best practices on page 17.

For additional details about Exchange 2010 database and logs performance factors, review Understanding Database and Log Performance Factors on Microsoft TechNet.

Table 4. 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

I/O size increase for Exchange Server 2010

Even though the number of user IOPS decreased for Exchange Server 2010 compared with previous versions of Exchange, the size of the I/O has significantly increased.

Between Exchange 2003 and 2010 there is an eight-fold increase in the size of the user I/O; thus, even though the number of I/Os per user has decreased, the sizes of the I/Os that the storage must handle have grown significantly.

Table 5. Exchange database transactional I/O sizes

Exchange 2003	Exchange 2007	Exchange 2010
4 KB random	8 KB random	32KB random 256 KB sequential (BDM read I/O)

From a storage perspective, the I/O for Exchange Server 2010 is significantly different from previous versions of Exchange. The figure below illustrates the average database transactional I/O size observed during Exchange Server 2010 Jetstress testing. Notice that the average read I/O is about 100 KB in this example, which is a result of the 32 KB database transactional I/O size mix with 256 KB BDM I/O size). We can expect the same results for a similar profile in a production environment as shown in Figure 1.

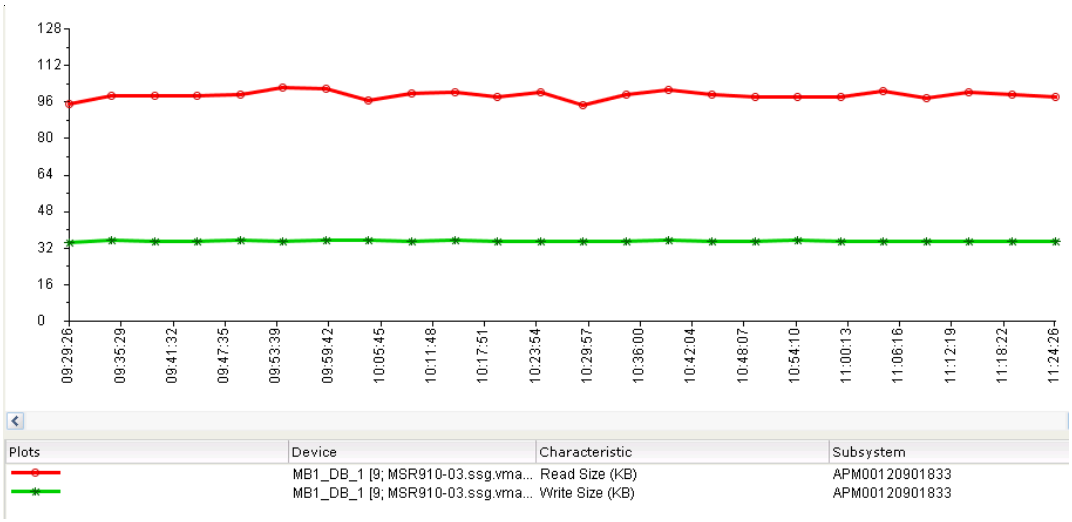

Figure 1. Average database transactional I/O size during Exchange Server 2010 Jetstress testing

Background database maintenance

One important difference in Exchange Server 2010 from the previous version is the way in which the database maintenance is performed. Now it includes online defragmentation and online database scanning. The process, called Background Database Maintenance (BDM), produces a large sequential 256 KB read I/O. BDM is enabled by default and runs 24x7 on both active and passive database copies.

With Exchange 2010 in a DAG configuration, you need to perform a consistency check only once a week, since BDM runs constantly and each shipped log is checked before it is applied to passive databases.

Important! You can schedule BDM processes to run against active copies, but you cannot schedule or disable them for passive DAG copies.

For more details about Exchange 2010 Background Database Maintenance process: http://blogs.technet.com/b/exchange/archive/2011/12/14/database-maintenance-in-exchange-2010.aspx

The I/O and throughput requirements for BDM need to be carefully considered as part of storage design when a large number of databases are deployed.

The large 256 KB BDM read I/O is factored into EMC's I/O calculation in the form of an additional EMC recommended 20 percent overhead factor, which is described in detail in the I/O calculations for number of disks required section of this document, which begins on page 35.

Bandwidth requirements for Background Database maintenance

BDM throughput (MB/s) requirements are especially significant if you deploy a large number of databases on a single bus or array. You must ensure sufficient throughput capability to permit BDM to run against all databases—both active and passive—simultaneously, without affecting performance.

EMC testing with Jetstress shows that BDM accounts for ~7.5 MB/s per database copy. The total throughput for each database is variable and depends on the additional throughput requirements of the user load. See the Bandwidth calculation section of this document, beginning on page 39, to determine the throughput requirements for each database in your design.

To illustrate the effects of BDM on throughput MB/s requirements, average total bandwidth consumption tests were run with and without BDM enabled. Figure 2 shows the database throughput for 500 users at 0.15 IOPS per user with BDM disabled, which averaged around 3.11 MB/s.

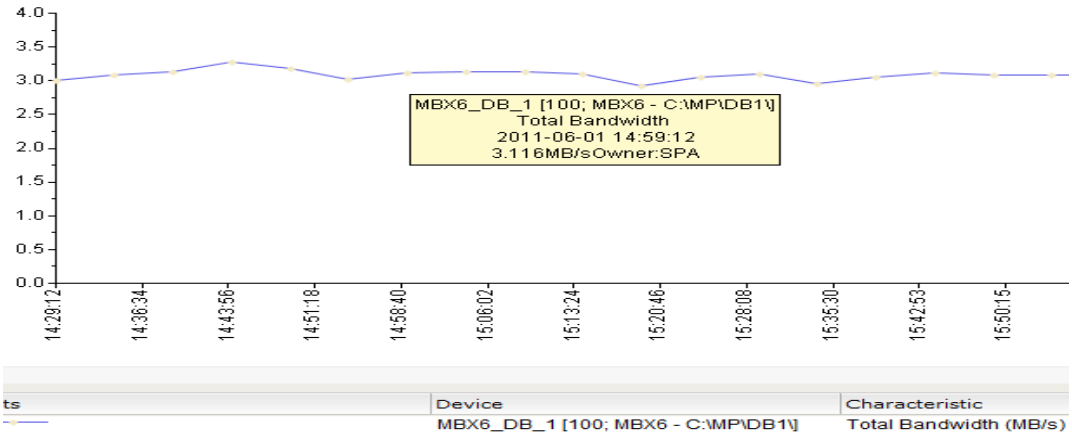

Figure 2. Database throughput for 500 0.15 IOPS users with BDM disabled

For the same database, with BDM enabled, the throughput increased to 10.4 MB/s, as shown in Figure 3. The ~7.5 MB/s difference was consistent across all databases and different user profiles. The effect of BDM on throughput can be significant when hundreds of databases are constrained to a single bus. Thus, it is necessary to consider the BDM effect when designing storage for Exchange Server 2010. We recommend deploying a lowering number of larger databases (500 GB to 2 TB), rather than a large number of smaller databases (up to 500 GB) to minimize bandwidth requirements.

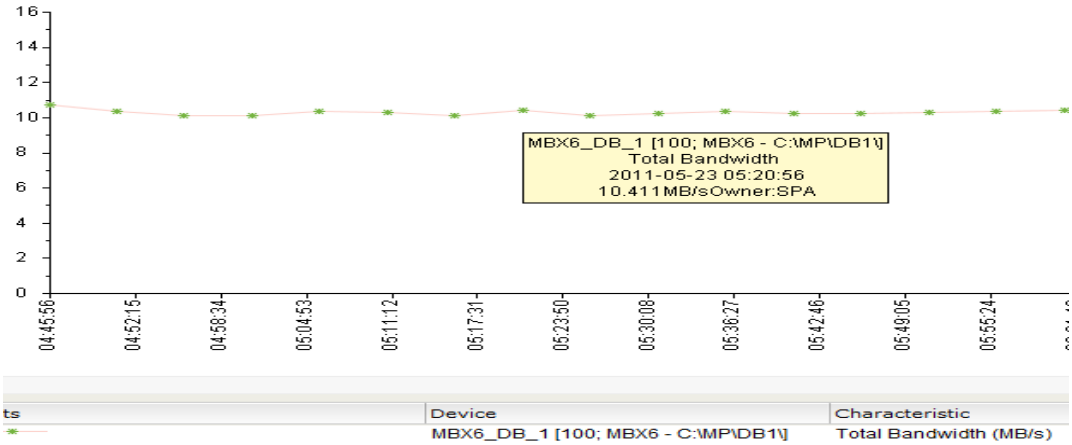

Figure 3. Total Bandwidth for 500 users per database with 0.15 IOPS user profile with BDM enabled

EMC storage design best practices

Storage design is one of the most important elements of a successful deployment of Microsoft Exchange Server 2010. To achieve a storage design that is optimal for reliability, performance, price, and ease of use, it is critical to follow the recommended guidelines provided by the storage vendor.

This section provides general best practices for deploying Exchange Server 2010 on EMC storage. This section also provides recommendations for deploying Exchange Server 2010 on EMC unified storage and EMC Symmetrix VMAX series storage, including recommendations for using specific EMC storage array features with Exchange Server 2010.

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 2010 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 2010 deployment depend on a number of factors, including your users' mailbox size limit and IOPS requirements.

EMC currently offers the following disk types with its VNX family of unified storage and Symmetrix VMAX series storage:

- **FC (10 and 15k rpm)** —Reliable disk drives with high read/write speeds. These disks are ideal for high I/O requirements but might not be appropriate for high capacity requirements.
- **SAS (10 and 15k rpm)** —An improvement of SCSI. SAS disks provide high capacity with moderate I/O speed, which makes them highly suitable for Exchange Server 2010 environments with high IOPS per user requirements.
- **SATA (VMAX) and NL SAS (VNX) (7.2k rpm)** – SATA and NL SAS disks are a good fit for the less demanding I/O requirements of Exchange Server 2010. These disks support large mailboxes at a relatively low cost. NL SAS disks are typically the best choice for larger mailbox sizes and average to heavy I/O profiles.
- **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 are not the best choice for Exchange data due to smaller capacity, but they can be appropriate when using thin provisioning on VNX as FAST Cache, or an automated storage tiering features such as EMC FAST VP or FAST Cache to handle any unanticipated I/O spikes.

Many IT organizations today are interested in deploying larger mailboxes for their customers. Since the Exchange Server 2010 storage engine now has lower I/O requirements, low-cost NL SAS and SATA drives are the 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 can be an appropriate choice, since these drives can handle high I/O requirements more efficiently.

As a general rule:

- 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.

IOPS per disk

Different disk types generate different numbers of Exchange Server 2010 IOPS. Consider this when calculating disk requirements for your environment. The following table provides large random disk IOPS data from the most recent Exchange validation on EMC VNX and VMAX storage. These results are subject to change based on future testing.

Important! EMC strongly recommends using values from Table 6 when calculating Exchange 2010 IOPS requirements for deployment on VNX and VMAX storage arrays.

Table 6. Exchange Server 2010 IOPS for various disk types on EMC storage

Disk type	Exchange Server 2010 IOPS per disk on VNX systems	Exchange Server 2010 IOPS per disk on VMAX systems
15K rpm FC/SAS	160	180
10K rpm FC/SAS	130	130
7.2K rpm SATA	N/A	55
7.2K rpm NL/SAS	65	N/A

RAID type selection

Selecting an appropriate RAID type for your environment is another important decision point for a successful implementation of Exchange Server 2010. Any RAID type can be used, provided there are enough disks to handle the I/O and storage capacity requirements. In general, RAID type decisions are based on customer requirements. In order to select an appropriate RAID type for your environment, consider your specific performance, capacity, and high availability requirements.

The EMC VNX family of unified storage and EMC Symmetrix VMAX series storage support RAID 1/0, RAID 5, and RAID 6 on Flash, FC, SAS, and SATA drives. 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. In general, RAID 1/0 is the best choice for Exchange Server 2010, especially if SATA and NL/SAS drives are used.
- 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 environments with low I/O requirements and where large mailboxes are being deployed.
- 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.

Table 7 shows RAID overhead, performance, and storage utilization information for each RAID type.

Note: The RAID overhead value becomes important when performing I/O calculations for number of disks required. For details, see the formula for applying RAID overhead on page 35.

Table 7. RAID type comparison

RAID type	RAID overhead value	Performance	Storage utilization
RAID 1/0 striping + mirroring	2	High	Low
RAID 5 striping + parity	4	Medium	High
RAID 6 striping + double parity	6	Low	Medium

Special considerations for VNXe storage

The EMC VNXe product family is designed around application-aware wizards for provisioning storage. These wizards automatically incorporate many of the best practices from this document, along with recommendations specific to the VNXe platform, into the storage design without additional user intervention. For a complete discussion of storage provisioning for Microsoft Exchange environments on VNXe storage, refer to *Microsoft Exchange 2010 or Exchange 2007 on EMC VNXe Series Deployment Guide.*

General best practices for Exchange Server 2010 on EMC storage

This section provides general storage design best practices that apply to the deployment of Exchange Server 2010 on EMC storage.

- To ensure the highest level of Exchange performance and predictability, isolate the Microsoft Exchange Server database and log workload from other I/O-intensive applications and workloads (assign the Exchange workload to its own set of disks). This also simplifies troubleshooting in the event of a storage-related performance issue. The single exception to this guideline is the use of properly sized Symmetrix FAST VP.
- When sizing storage, always calculate disk I/O requirements before calculating capacity requirements and select the appropriate disk types that meet your I/O and capacity requirements. Use the Exchange storage building block design approach whenever possible (see the section entitled Exchange storage building block on page30.
- If your storage design does *not* include replication, do *not* place both database and log files for the same database on the same physical disks (otherwise neither the database files nor the log files are adequately protected against disk failure). You may, however, place the database files from one database and the log files from another database on the same physical disks.
- If your storage design includes replication, you can consider placing 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.
- If DAGs are used, the database can be as large as 2 TB. Select a database size that is appropriate for your specific requirements. Consider backup and restore times when choosing a database size.
- Although 2 TB databases size are supported, ensure that the hypervisor support for LUNs larger than 2 TB if you are using virtualization. At the same time, consider a 20 percent free space buffer and other data, such as an index, that resides on the same volume as the database. Note, that the 2 TB database cannot fit on a 2 TB LUN.
- To minimize BDM impact, consider deploying a lower number of larger databases (500 GB to 2 TB), rather than a larger number of smaller databases.
- Deploy each DAG copy on its own set of physical disks. For example, do not put multiple dag database copies into the same RAID group or pool.
- Spread the load as evenly as possible across storage array resources (VMAX engines, VNX storage processors, and so on).
- Format Windows NTFS volumes used for Exchange databases and logs with an allocation unit size of 64 KB.
- Use the Exchange storage building block design approach whenever possible (see Exchange storage building block on page 30).

FAST VP with Exchange 2010

EMC Fully Automated Storage Tiering for Virtual Pools (FAST VP) automatically optimizes performance in a tiered environment reducing overall costs, footprint and management effort. FAST VP puts the right data in the right place at the right time and is a game changing technology that delivers automation and efficiency to the virtual data center. FAST VP 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.
- Customers 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.
- Customers can choose to have data placed on the highest or lowest available tier, ensuring that performance and cost commitments are met.
- Customers can define data movement schedules to minimize FAST management responsibilities.

Whether to use FAST VP with Exchange Server 2010 on VNX or VMAX storage arrays depends entirely on customer requirements. If the customer is moving toward the use of FAST VP for all applications, or if the customer needs to be able to handle unanticipated spikes in performance demand, FAST VP might be a good fit. In most situations FAST VP is not the lowest cost option from a storage perspective, but the higher initial cost can be reclaimed through reduced administration and faster space and I/O issue resolution.

Due to the changes in the Exchange 2010 storage architecture resulting in lower I/O to storage devices and the trend to deploy larger mailboxes, many Exchange designs are capable of utilizing high capacity low RPM drives (for example, 7.2k rpm SATA/NL SAS). There are, however, Exchange configurations with considerably higher I/O demands and smaller mailbox requirements that would benefit from adding SAS/FC and even a FLASH tier into an existing storage pool created with NL-SAS or SATA drives. This will eliminate spikes in application performance and allow customers who wish to use larger less expensive NL-SAS/SATA drives, but require the speed from faster SAS/FC or FLASH drives.

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, review *EMC FAST VP for Unified Storage Systems* white paper available on www.emc.com, which describes Fully Automated Storage Tiering for Virtual Pools (FAST VP) technology, its features, and implementation guidelines.

For EMC VMAX, review *EMC Symmetrix VMAX with Enginuity* white paper available on www.emc.com.

VNX-specific storage design guidelines

The EMC VNX family of unified storage offers various features and capabilities including FAST Cache, storage pools, and thin LUNs. It is important to understand which of these features are compatible with and beneficial for Exchange Server 2010. This section presents answers to some typical questions about the deployment of Exchange Server 2010 on VNX storage.

VNX settings

What are the most important VNX settings for Exchange 2010?

- Set the storage array page size parameter to 16 KB.
- The system memory setting depends on the array model. Allocate 1 GB to read cache and the rest of the memory to write cache.
- Enable both read and write cache.

Pools or RAID Groups with Exchange 2010 on VNX

Should I use storage pools or RAID groups with Exchange Server 2010 on VNX?

You can use either RAID groups or storage pools for Exchange Server 2010 on VNX. The use of storage pools simplifies storage provisioning. Traditional storage provisioning with RAID groups only 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. Such pool-based provisioning provides benefits similar to metaLUNs striping across many drives but, unlike metaLUNs, storage pools require minimal planning and management effort. Storage pools support the same RAID protection levels as RAID groups do: RAID 5, RAID 6, and RAID 1/0.

You are required to use storage pools when LUNs are thin provisioned or when VNX Snapshots are used for Exchange protection. When using storage pools, the following guidelines apply:

- Use homogeneous storage pools (pools with the same disk type) for Exchange 2010 data, although the use of FAST VP, multi-tiered pools is acceptable.
- Design and grow storage pools by using the appropriate multiplier for best performance (R1/0 4+4, R5 4+1, R6 6+2).
- When you use VNX Snapshots, you must configure Exchange storage (databases and logs) in storage pools.
- When using thin pools, consider using a thin pool utilization thresholds tool to monitor the pools to prevent thin pools from running out of space.
- When using RAID groups with RAID 5, do not stripe metaLUNs across multiple RAID groups. Doing so can reduce performance, since Exchange Server 2010 is a high-I/O and high-bandwidth application.

Thin LUNS with Exchange 2010 on VNX

Can I use thin LUNs with Exchange Server 2010 on VNX?

Thin provisioning can provide significant savings when large mailboxes are deployed. Use these guidelines for deploying Exchange 2010 on thin LUNs:

- Thin LUNs can be deployed for lower to medium Exchange 2010 user workloads.
- If you are using FAST VP pools with slow speed capacity drives (NL SAS) and other higher speed drives (10k, 15k SAS or FLASH), we recommend that you set the tier policy settings to **Highest Available Tier** first, then **Auto-Tier**. Always use R10 4+4 for the faster drive tiers.
- Although it is not required, you should enable FAST Cache on pools with Thin LUNs. The exception where FAST Cache is required is when the LUN is protected with VNX Snapshots.
- Use of RAID 10 is recommended for Thin LUNs

FAST Cache with Exchange 2010 on VNX

Should I use FAST Cache with Exchange Server 2010 on VNX?

Due to lower Exchange 2010 I/O user profile requirements, there is rarely a need for any additional cache boost for most configurations on VNX systems. Exchange 2010 utilizes considerable host cache and coalescing of data, which works counter to FAST Cache technology. However, the use of FAST Cache on VNX will benefit Exchange 2010 performance in certain configurations. Specifically, VNX FAST Cache is required or recommended in the following configurations:

- Strongly recommended when databases LUNs are thin provisioned
- Required when you use VNX snapshots for protection
- Required on thin or thick LUNs when VNX snapshots are enabled

Follow these guidelines when using FAST Cache with Exchange 2010:

- If you use FAST Cache with Exchange Server 2010, segregate the database volumes from the log volumes by placing each volume type in a separate storage pool, RAID group, or on different LUNs. Enable FAST Cache only on database LUNs. Do not enable FAST Cache on log LUNs.
- We do not recommend that you use FAST Cache for logs, even if you are using VNX Snapshots.
- Do not use FAST Cache for the purpose of making a storage system pass Jetstress tests.
- Enable FAST Cache on storage pools housing Exchange databases, with thick or thin LUNs, that are configured for VNX Snapshots. As a general sizing recommendation, use a 1000:1 Exchange data to FAST Cache ratio. That is, for 1,000 GB of Exchange data, use approximately 1 GB of FAST Cache.

FAST VP with Exchange 2010 on VNX

Should I use FAST VP with Exchange Server 2010 on VNX?

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 2010 on VNX, follow these guidelines:

- When using FAST VP, we do not recommend placing database files and log files in the same pool, since log files have a low I/O requirement and do not need to be moved to a higher tier.
- Always place Log files in separate storage pools, and always using RAID10.

  Using RAID1/0 and separating log files dramatically reduces the total I/O to disk. This can greatly reduce VNX Storage Processor (SP), BUS, and DISK utilization due to the methods in which VNX FLARE Operating environment and Write Cache handles small sequential I/Os.
- 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 Policy for the participating pool LUNs to the Highest Available Tier first, then Auto-Tier.

SOAP tool

SOAP Tool is a utility that optimizes VNX Storage Pool Thick LUNs for maximum performance after the LUN creation and prior to disk partitioning on the Exchange server. You should run this utility for customers requiring uniform, deterministic high performance and low latencies across all LUNs within a VNX storage pool. We recommend that you use this tool when deploying Exchange in storage pools only on CLARiiON CX4 and VNX systems with FLARE release prior to FLARE 32. For more details about this tool visit http://powerlink.emc.com

Symmetrix VMAX-specific storage design guidelines

The following list includes several important design considerations for Exchange Server 2010 on VMAX:

- When creating LUNs, use fewer but larger hyper volumes to improve performance.
- Use a minimum of two HBAs per server, with each HBA connected to at least two director ports (across multiple VMAX engines, if possible).
- When using regular (Thick) LUNs, use striped meta volumes.
- When using Thin LUNs, use concatenated meta volumes.
- Keep logs out of the FAST VP policy by placing log files in a separate storage group from database files.

As with the EMC VNX family of unified storage, a number of popular questions arise around some of the key EMC Symmetrix VMAX series features as they relate to Exchange Server 2010.

Thin LUNs with Exchange 2010 on VMAX

Should I use thin LUN pools with Exchange Server 2010 on VMAX?

Thin LUN pools are recommended for Exchange Server 2010 on Symmetrix VMAX. Thin device performance is equivalent to regular (thick) device performance on VMAX, and in most cases, the use of thin pools can reduce the initial storage requirement.

If you use thin pools, you must ensure that the initial configuration of disks can support the I/O requirements. A thin pool can be configured to support a single Exchange building block or multiple building blocks, depending on customer requirements.

When using thin pools, consider using the thin pool utilization threshold tool to monitor the pools and thus avoid thin pools running out of space.

Exchange IOPS per VMAX engine

How many Exchange Server 2010 user IOPS can a VMAX engine support?

The number of Exchange Server 2010 user IOPS that a single VMAX engine can support depends on multiple factors including the amount of cache in the array, the disk type (SATA or FC), the RAID type, FA and DA utilization, and the number of databases.

FAST VP configuration guidelines with Exchange 2010 on VMAX

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 2010 on VMAX, follow these guidelines:

- When using FAST VP, we do *not* recommend that you place database files and log files in the same pool because log files have a low I/O requirement and do not need to be moved to a higher tier.
- Keep logs out of the FAST VP policy by placing the log files in a separate storage group from database files.

- When using FAST VP with DAG, do not place DAG copies of the same database in the same pool on the same disks.
- Perform sizing according to FAST VP policy requirements. A good starting point is an 80/20 skew assumption.

  For example, perform the appropriate calculations to support 80 percent of I/O on Fibre Channel disks and 20 percent of I/O on SATA disks.
- Reverse the skew for disk space: Place 80 percent of the data on SATA disks and 20 of the data on Fibre Channel disks.
- Perform the sizing exercise only for Fibre Channel and SATA tiers, and allocate a relatively small amount of Flash disk space to handle unanticipated spikes.

  Note: The skew ratios can vary and depend on the specific Exchange profile.

VFCache with Exchange 2010

EMC VFCache is a server Flash caching solution that reduces latency and increases throughput to dramatically improve application performance by leveraging intelligent software and PCIe Flash technology. VFCache accelerates reads and protects data by using a write-through cache to the networked storage to deliver persistent high availability and disaster recovery. VFCache 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.

VFCache benefits

VFCache provides complete and flexible control over the scope and granularity at which it can be enabled. In physical environments, you can enable or disable VFCache at the source volume level or LUN level. In virtual environments, the VFCache capacity is provisioned to an individual virtual machine (VM). The allocated cache capacity inside the VM is then configured at the virtual disk level.

To learn more about EMC VFCache review the following documentation on http://www.emc.com/storage/vfcache/vfcache.htm and http://www.emc.com/collateral/hardware/data-sheet/h9581-vfcache-ds.pdf

How does VFCache accelerate Exchange performance?

When VFCache is configured for volumes housing Exchange databases, VFCache accelerates block I/O reads that require the highest input/output operations per second (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 in the server on the PCIe bus, VFCache bypasses the overhead of network storage access, thus reducing response times.

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

Note: Performance gain and reduction in response times will vary based on each customer's Exchange email usage. We highly 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.

For additional details about VFCache performance with Exchange 2010, review white paper titled *Accelerating Microsoft Exchange 2010 Performance with EMC VFCache* at www.emc.com.

Virtualization best practices for Exchange Server 2010 on EMC storage

In recent years, virtualizing most mission critical applications as well as the Microsoft Exchange Server 2010 environment has become a popular choice for many companies because of the many benefits virtualization can provide, including improved ROI, ease of management, greater HA and DR capabilities, and flexibility. With the newer, more powerful servers, you can achieve a six-to-one server consolidation ratio, which can yield significant cost savings. The virtualization of an Exchange Server 2010 environment requires some unique storage design best practices.

The Windows Server Virtualization Validation Program (SVVP), available on the Microsoft website (http://www.windowsservercatalog.com/svvp.aspx), provides information about the ways in which EMC storage supports the virtualization of Exchange Server 2010 environments.

By virtualizing an Exchange Server 2010 environment hosted on EMC VNX family or Symmetrix VMAX series storage, customers can utilize features such as VMware® VMotion™ and Microsoft Hyper-V live migration tools. These features enable virtual servers to be moved between different server hardware without application disruption.

General guidelines for Exchange virtualization

Some general guidelines that apply to the virtualization of Microsoft Exchange Server 2010 are presented below.

- Core Exchange design principles still apply:
  - Design for performance, reliability and capacity.
  - Design for user profiles (CAS/MBX roles).
  - Design for message profiles (Hub/Edge).
- Size virtual machines according to the Exchange role (MBX, CAS, Hub, Multi-role).
- Physical sizing still applies.
- Size virtual machine root servers to accommodate the number of guests that need to be supported.
- Spread DAG copies across multiple physical hosts to minimize potential downtime in the event of physical server issues.
- For Microsoft Hyper-V, add 10 percent to CPU requirements for hypervisor overhead. For VMware vSphere, add about five percent.
- Exchange Server virtual machines, including Exchange mailbox virtual machines that are part of a Database Availability Group (DAG), can be combined with host-based failover clustering and migration technology, provided the virtual machines are configured so that they do not save and restore state on disk when moved or taken offline.
- Following the failover (or failback) of a virtual machine from source to target, the VM must be restarted (cold started) on the target.

- PowerPath or PowerPath/VE is highly recommended to be installed on physical Hyper-V or ESX hosts for load balancing, path management, and I/O path failure detection.
- Disable migration technologies that save state and migrate. Always migrate live or completely shut down VMs.
- Dedicate and reserve CPU and memory resources for the mailbox VMs; do not over-commit these resources.
- Disable hypervisor-based auto tuning features—dynamic memory, storage tuning and rebalancing.
- Ensure that each server has at least four paths (two HBAs) to the storage, with four ports total.
- Place Exchange storage on separate disks from guest OS (VHD/VMFS or VMDK) physical storage.
- Know the hypervisor limits, for example: 256 SCSI disks per host, processor limits, and memory limits.
- Use both hypervisor and OS performance tools to monitor the VMs.
- Use appropriate volume types. For VMware vSphere, use VMFS, RDM, or iSCSI on the host or guest. For Microsoft Hyper-V, use an SCSI pass-through device, VHD, or iSCSI on the host or guest.
- Configuring iSCSI storage to use an iSCSI initiator inside an Exchange guest virtual machine is supported. However, there will be reduced performance in this configuration if the network stack inside a virtual machine isn't full-featured (for example, not all virtual network stacks support jumbo frames).

Unsupported configurations

Microsoft does not support the following configurations:

- Hypervisor Snaps
- › 2:1 vCPU to pCPU ratio
- Applications running at the root (Hyper-V only)
- VSS backup of root for pass-through disks

Resources for Exchange virtualization

Notes

- Review and leverage the recommendations from the document entitled *Microsoft Exchange 2010 on VMware Best Practices Guide* available at http://www.vmware.com/files/pdf/exchange-2010-on-vmware-best-practices-guide.pdf when deploying Exchange on a vSphere platform.
- Review Microsoft's guidelines entitled *Understanding Exchange 2010 Virtualization* at http://technet.microsoft.com/en-us/library/jj126252.aspx for additional information and recommendations for virtualizing Exchange Server 2010 services on a Hyper-V platform.

Exchange storage building block

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.

Sizing and configuring storage for use with Exchange Server 2010 can be a complicated process, driven by many variables and requirements, which vary from organization to organization. Properly configured Exchange storage, combined with optimally sized server and network infrastructures, can guarantee smooth Exchange operation and 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 2010 is to define a unit of measure—a *mailbox server building block*.

What is a 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 2010 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 2010.

Once the initial building block is designed, it can be easily reproduced to support the required number of users in your enterprise. By using this approach, EMC customers can now 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 4 illustrates the Exchange building block design process.

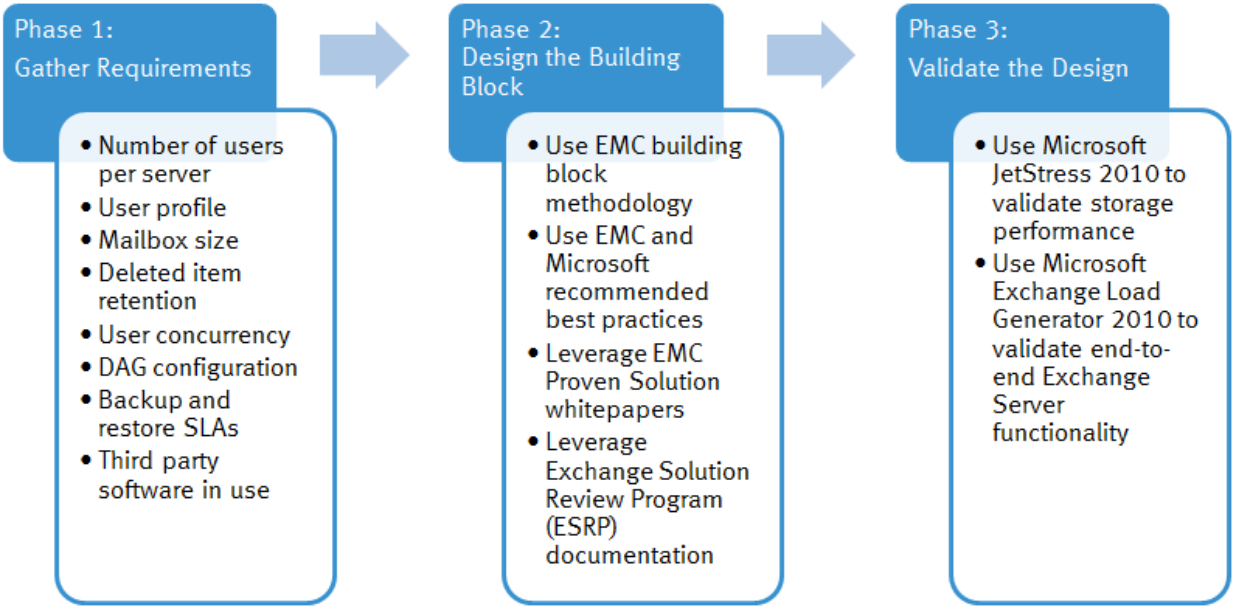

Figure 4. Phases in Exchange building block design

The sections that follow present examples of designing an Exchange Server 2010 building block on EMC storage by using a three-phase approach. The focus is on the storage aspect of designing an Exchange Server 2010 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 state that they must be agreed upon by all parties. If the requirements change it is essential that all relevant parties are informed of the changes and an assessment is made of the changes' potential impact on the project.

Key Exchange Server 2010 requirements

Key requirements for Exchange Server 2010 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 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)

Example customer requirements

The following table presents an example set of Exchange Server 2010 customer requirements.

Table 8. Example Exchange Server 2010 customer requirements

Item	Value
Total number of users (mailboxes) in Exchange environment	16,000
Mailbox size limit	1.5 GB (~1,500 MB) per user
Number of messages sent and received per user per day	150 messages per user per day (0.15 IOPS)
Target average message size	75 KB (~.75 MB)
Outlook mode	Cached mode, 100 percent MAPI
Number of mailbox servers	8
Number of mailboxes per server	4,000 in a switchover configuration (2,000 active/2,000 passive)
HA requirements	One DAG, two database copies per DAG
Number of databases per server	8
Number of users per database	500
Deleted items retention (DIR) period	14 days
Log protection buffer (to protect against log truncation failure), in days	3 days
24 x 7 BDM configuration	Enabled
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)	n/a

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 **Exchange 2010 Mailbox Server Role Requirements Calculator** tool from Microsoft provides CPU and memory guidance's in addition to storage recommendations. EMC enhances Microsoft's calculator and adds special EMC extensions, 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 on EMC storage. Specifically, the EMC calculator includes thin provisioning features for sizing Exchange storage.

For access to the *Microsoft Exchange 2010 Mailbox server Role Requirements Calculator*, visit the Microsoft Exchange Team Blog at http://msexchangeteam.com/archive/2009/11/09/453117.aspx.

For access to the *Microsoft Exchange 2010 Mailbox server Role Requirements Calculator with EMC extensions*, visit "Everything Microsoft at EMC" section on EMC Community Network portal at https://community.emc.com/docs/DOC-13037.

Make sure to consult EMC for any additional guidelines during the design and deployment phases.

While the Microsoft and EMC calculator 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. This section presents an example of how to calculate space requirements. The process is based on information provided on TechNet at http://technet.microsoft.com/en-us/library/ee832789.aspx.

I/O calculations for number of disks required

Phase two in the building-block design phase involves defining a building block that satisfies the requirements collected in phase 1. The building block includes storage, server, and network resources. This document focuses on defining the storage resources.

I/O calculation

The first step in determining the storage resources required for your building block is to calculate disk requirements based on I/O. The sequence of relevant formulas follows.

User IOPS

First, calculate user IOPS. User IOPS is the transactional database IOPS and also the target for Jetstress testing. Here is the formula, with data value descriptors enclosed in angle brackets:

User IOPS = ‹number of users› * ‹IOPS per user› + ‹any additional IOPS overhead›

Here is the formula again, this time including sample values and a calculated solution (angle-bracketed descriptors now shown in grey):

User IOPS = ‹number of users› **4,000 *** ‹IOPS per user› **0.15 +**
‹any additional IOPS overhead› **0 = 600**

Note: Microsoft typically recommends a 20 percent I/O overhead factor to account for special customer requirements and future growth. A value of zero is used in this example.

Front-end IOPS

Next, add an EMC required 20 percent of the user IOPS value to account for logs, BDM, and so on. This yields the total front-end Exchange Server 2010 IOPS.

EMC 20% overhead

Front-end IOPS = ‹user IOPS› + (‹user IOPS› * 0.20)

Front-end IOPS = 600 + 120 = 720

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

RAID overhead and disk I/O

Next, apply the appropriate RAID overhead value (from Table 7 on page 19) to the write I/O, and then divide the result by the appropriate Exchange Server 2010 IOPS per disk value from Table 6 on page 18.

For example, with a read/write ratio of 3:2 (60/40 percent split) and VNX NL/SAS drives, the formula would be:

----------60% reads---------- *----------40% writes----------*

((‹front-end IOPS› * 0.60) + ‹RAID overhead value› * (‹front-end IOPS› * 0.40)) / ‹IOPS per disk› = ‹total number of disks required per building block›

RAID type choices

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 space. Table 9 presents the equation with various RAID types using NL-SAS disks on VNX storage. Based on this data, RAID 1/0 is the best choice from an I/O perspective.

Table 9. Calculations for using NL_SAS disks on VNX storage

RAID type	Number of disks required
With RAID 1/0 (overhead value of 2)	(720 * 0.60) + 2(720 * 0.40) = 1,008 / 65 = 16 disks
With RAID 5 (overhead value of 4)	(720 * 0.60) + 4(720 * 0.40) = 1,584 / 65 = 26 disks
With RAID 6 (overhead value of 6)	(720 * 0.60) + 6(720 * 0.40) = 2,160 / 65 = 34 disks

Capacity calculations for disks requirements

If you are designing Exchange 2010 solution to leverage thin provisioning, you need to perform two sets of calculations for determining the disk capacity requirements:

- The first calculations determine the initial capacity requirements.- This is necessary to identify the storage requirements to support the initial mailbox capacity. You will base the storage capacity purchase on these calculations.
- The second calculations determine 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 2010 and leveraging thin provisioning on EMC storage, it is a best practice to separate the Exchange transactional logs from the database thin 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. You do not need to thin provision the Exchange log LUNs.

Database volume size calculation

In order to determine database volume size requirements, you must first calculate:

- User mailbox size on disk
- Database size on disk
- Database LUN size

User mailbox size on disk formula

The formula for user mailbox size on disk is:

User mailbox size on disk = ‹mailbox size› + ‹whitespace› + ‹dumpster›

White space formula

White space = ‹number of messages sent/received per day› * ‹average message size›

Example: White space = 150 messages per day * 75 KB / 1,024 MB = ~11 MB

Dumpster formula

Dumpster = (‹number of messages sent/received per day› * ‹average message size› * ‹number of days for deleted items retention›) + (‹mailbox size› * 0.012) + (‹mailbox size› * 0.03)

Example: Dumpster = (150 messages per day * 75 KB / 1,024 MB * 14) + (1,500 MB * 0.012) + (1,500 MB * 0.03) = 217 MB

Once the whitespace and dumpster values are determined, prepare a table such as this to help you determine the database size on disk and database LUN size:

Mailbox size	White space	Dumpster	User mailbox size on disk	Number of mailboxes per database
1,500 MB	11 MB	217 MB	1,728 MB	500 users

Example: User mailbox size on disk = 1,500 MB + 11 MB + 217 MB = 1,728 MB

Database size on disk formula

Database size on disk = ‹number of mailboxes per database› * ‹mailbox size on disk› * ‹database overhead growth factor›

- Database overhead growth factor is 1.2 (effectively adds 20 percent)

Example: Database size on disk = 500 users * 1,728 MB * 1.2 = 1,037 GB

Database LUN size formula

Database LUN size = (‹database size on disk› + ‹content index factor›) / (1 - free space percentage requirement factor)

- The content Index is usually 10 percent of the database size. So the content index factor is 1.1
- Database LUN recommended 20percent free space requirement factor is 0.2

Example: Database LUN size for fully provisioned mailboxes = 1,037 GB * 1.1 = 1140.7 / 0.8 = 1,426 GB

For more information on the Microsoft formula for database LUN size, visit http://technet.microsoft.com/en-us/library/ee832789.aspx.

Log LUN size

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.
- **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, 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 on a nightly or weekly basis 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, it is recommended to 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.

Log LUN size formula

Log LUN size = ‹logs per day, per mailbox› * ‹number of mailboxes per database› * ‹protection buffer from log truncation failure, in days› * ‹free space requirement factor›

- Free space requirement factor is 1.2 (effectively adds 20 percent)

Example: Log LUN size = 30 logs at 1 MB per log * 500 * 3 * 1.2 = 54 GB

Total space required per building block

Table 10 presents the example database and log LUN size requirements and the resulting total space required per building block.

Table 10. Space requirements per building block

Total database LUN size per server	Total log LUN size per server	Total LUN size capacity required per server
11,408 GB (1,426 GB x 8 LUNs per server)	432 GB (54 GB x 8 LUNs per server)	11,840 GB (16 LUNs)

Total space required per building block formula

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

Example: Total LUN space required per server = (1,426 + 54 GB) * 8 = 11,840 GB

Total number of disks required

Example: Total number of disks required = 11,840 GB / 1,834 = ~8 disks

Note: The number of disks is "rounded up" to eight. 1,834 GB represents the usable capacity of 2 TB NL-SAS disks on VNX systems.

With RAID 1/0, the required number of 2 TB NL-SAS disks becomes 16.

Summary of disk requirements

The best configuration, based on the example requirements and calculated values, is 16 2 TB NL-SAS disks. This configuration meets both IOPS and capacity requirements for a building block of 4,000 users with a 0.15 IOPS user profile and a 1.5 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 Exchange environment. Following the example:

Total number of disks required 128 disks (16 * 8) to support 16,000 users with 1.5 GB mailbox quota in a two-copy DAG.

Bandwidth calculation

If more than 100 Exchange databases are to be deployed on a single array, it is recommended to perform I/O throughput validation to ensure that the array buses do not become an I/O bottleneck.

The throughput validation process involves multiple steps:

1. Determine how many databases the customer requires
2. Ensure that the database LUNs are evenly distributed among the backend buses and storage processors
3. Determine the throughput MB/s per database
4. Determine the required throughput MB/s per bus
5. Determine whether each bus can accommodate the peak Exchange database throughput

Throughput MB/s per database

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

Throughput MB/s per database = ‹total transactional IOPS per database› * 32 KB + 7.5 MB

Where:

- 32 KB is a page size
- 7.5 MB is an estimated BDM throughput per database

Required throughput MB/s per bus

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 MB/s per bus.

Required throughput MB/s per bus = ‹throughput MB/s per database› * ‹total number of active and passive databases per bus›

Important! The number of databases per bus is the total number of active and passive databases per bus.

Compare your throughput requirement with the array's bus capability

Next you must compare your 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. Contact your local speed guru to obtain the bus specifications for the specific array you are using, since these specifications change over time, or use *Disksizer* for CX4 or VNX.

Example calculations

Assumption: 500 at 0.15 IOP users per database and 200 databases per bus

Total transactional IOPS per database = (500 * 0.15) * 32 KB = 2.4 MB/s

Throughput MB/s per database = 2.4 MB/s + 7.5 MB/s = 9.9 MB/s

Required throughput MB/s per bus = 9.9 MB/s * 200 databases = 1,980 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.

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 the metrics defined in that profile. It is strongly recommended that Jetstress testing is used to validate storage reliability and performance prior to the deployment of the Microsoft Exchange production environment.

Additional information about Jetstress is available at this URL:

http://technet.microsoft.com/en-us/library/dd335108.aspx

For comprehensive information on using Jetstress, visit this URL:

http://technet.microsoft.com/en-us/library/ff706601.aspx

ESRP reports

Another resource to use is a Microsoft program for validating storage vendors' Exchange Server 2010 designs known as the Exchange Solution Review Program (ESRP). 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. 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 2010 storage design are available at http://www.emc.com/exchange and http://www.microsoft.com/emc.

High Availability and Data protection for Exchange Server 2010 on EMC storage

In Exchange 2010 Microsoft enhanced its native Exchange high availability and data protection capabilities by introducing the Database Availability Group (DAG) feature.

Exchange 2010 DAG

A DAG is a set of Microsoft Exchange Server 2010 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 you add a Mailbox server 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. Various Exchange server 2010 high availability and disaster recovery options are described in this section.

Lagged copy

If a past point-in-time copy of mailbox data is a requirement, Exchange provides the ability to create a lagged copy in a DAG environment. This can be useful in the rare event that there's a logical corruption that replicates across the databases in the DAG, resulting in a need to return to a previous point in time. It may also be useful if an administrator accidentally deletes mailboxes or user data. But lagged copies cannot provide multiple point-in-time recoveries. After using a lagged copy for recovery, you must create new lagged copy. This process may be resource and time consuming in some large environments. For additional information about Exchange database corruption types and lagged copy limitations, visit http://technet.microsoft.com/en-us/library/dd335158.aspx.

EMC has the ability to provide the same or better protection levels but using far less storage with the use of snapshots.

EMC Replication Enabler for Exchange 2010

EMC Replication Enabler for Microsoft Exchange Server 2010 (REE) is a free plug-in that integrates with Microsoft Exchange to enable storage-based replication to be used in place of DAG based replication to provide high availability and disaster recovery. This solution provides significant storage saving over using native DAG replication. For more information about REE, visit EMC's online support website at http://www.emc.com/replication/replication-enabler-exchange-2010.htm. For an EMC proven solution leveraging REE, review white paper entitled *Zero Data loss Recovery for Microsoft Exchange Server 2010* available on www.emc.com.

EMC high availability and data protection offerings for Exchange Server 2010

While Exchange 2010 native data protection is sufficient for some customers, many others still require full backup and restore capabilities for Exchange 2010 databases. EMC offers several options to provide high availability and data protection with Exchange 2010, including:

- EMC AppSync (with VNX Snapshots or RecoverPoint)
- EMC Replication Manager (EMC SnapView™ and EMC TimeFinder®)

- EMC Avamar®
- EMC Networker®

Each above mentioned product has its own strengths and capabilities. The decision about which one to choose depends on the individual customer's requirements for backup and restore and what type of EMC storage array they currently have or planning to purchase.

EMC hardware-based snapshot and clone products have been integrated with Microsoft VSS technology for many years. Symmetrix 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 Symmetrix systems. Replication Manager coordinates the entire data replication process—from discovery and configuration to the management of multiple, application-consistent, disk-based replicas. You can use the *Autodiscover* feature to discover the replication environment and enable streamlined management by scheduling, recording, and cataloging replica information including auto-expiration.

Visit www.emc.com on the Replication Manager product page for additional details.

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 granularity. AppSync also provides an application protection monitoring service that generates alerts when the SLAs are not met.

Visit www.emc.com on the AppSync product page for additional details.

For a complete solution leveraging Exchange 2010 with EMC AppSync and VNX snapshots, review white paper entitled *Advanced Protection for Microsoft Exchange2010 on EMC VNX Storage* at www.emc.com

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 2010 DAG feature is to be used, but without a lagged copy, EMC strongly recommends an alternative, solid, point-in-time Exchange data protection strategy to guard against logical corruption events.

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 2010 DAG feature is to be 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.

Disaster recovery options for Exchange 2010

EMC offers various DR options for Exchange Server 2010. Table 11 presents the three most popular options in addition to the native Exchange DAG feature. Each option has its advantages and disadvantages. The option that is best for you is determined by your specific DR requirements.

Table 11. EMC disaster recovery offerings for Exchange 2010

Offering	Replication method	Description
Cross-site (stretch) DAG	Exchange continuous replication (DAG)	Built into Exchange 2010 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.
Replication Enabler for Exchange (REE)	EMC RecoverPoint	Synchronous solution. Integrates with Exchange 2010 third party replication API. Replaces native DAG replication.

Additional backup recommendations

For Exchange 2010 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.
- Schedule backups to take place during off hours, whenever possible.

Conclusion

This document highlights the key decision points in planning a Microsoft Exchange Server 2010 deployment with EMC storage systems. Multiple configuration options are available to suit most requirements for any customer. EMC storage and data management software products are designed to provide customers the flexibility to manage their Exchange Server 2010 environments in a manner that best meets their business needs.

Best practices for designing Exchange Server 2010 storage are constantly evolving. This document presents a snapshot of the current best practices recommended by EMC for deploying Exchange Server 2010 with the EMC VNX family of unified storage or EMC Symmetrix VMAX series storage. Following these guidelines can greatly assist you in achieving an efficient, high-performance, and highly available Exchange Server 2010 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 2010
- Apply VNX- and VMAX-specific best practices for Exchange Server 2010
- Utilize an Exchange Server 2010 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 2010

Additional information

For additional guidance on deploying Microsoft Exchange Server 2010 with EMC VNX family of unified storage or EMC Symmetrix VMAX series storage, consult your local EMC Microsoft Exchange expert.