White Paper

MICROSOFT EXCHANGE 2010: STORAGE BEST PRACTICES AND DESIGN GUIDANCE 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.

November 2011

Microsoft GOLD CERTIFIED Partner

EMC²

Part Number h8888.2

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.
- An additional difference is reduced IO, made possible by a new database storage schema, which increases the database IO 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

The following terms are used in this document.

Exchange database availability group (DAG)—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—Exchange Server 2010 database maintenance now includes online defragmentation and online database scanning. This process, called Background Database Maintenance (BDM), produces a large sequential 256 KB IO 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 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 logical device that is configured and presented to the host but that allocates physical storage only as needed. The physical storage used to supply disk space to the thin devices comes from the virtual pool to which the thin device is bound. Thin devices can be created with an inflated capacity, because the actual storage space for data written to them is provided by data devices. To a host operating system, thin devices look like standard devices with their configured capacity, and the host interacts with them in the same way as standard devices.

FAST VP—Fully Automated Storage Tiering with Virtual Pools. FAST VP now supports sub-LUN automated tiering for file systems. FAST VP offers the following cost and performance benefits to customers:

- Customers can set policies to automatically tier data based on IO, avoiding the pre-provisioning tasks of determining on which tiers 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.
- FAST can run on two or more drive types, optimizing an investment in Fibre Channel, SATA, Flash, SAS (performance), and/or NL-SAS (capacity) drives.

FAST Cache—FAST Cache software 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. These additional configurations are only offered on the VNX.

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.

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.

- This suite, only available for the VNX series, includes Fully Automated Storage Tiering with Virtual Pools (FAST VP), FAST Cache, Unisphere Analyzer, and Unisphere Quality of Service Manager.

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.

- This suite for VNX series includes Event Enabler (anti-virus, quota management, auditing), File-Level Retention, and Host Encryption.
- This suite for VNXe series includes Event Enabler (anti-virus) and File-Level Retention.

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.

- This suite for VNX series includes SnapView™, SnapSure™, and RecoverPoint/SE CDP (continuous data protection).
- This suite for VNXe® series includes SnapSure.

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.

- This suite for VNX series includes Replicator, MirrorView/A, MirrorView/S, and RecoverPoint/SE CRR (continuous remote replication).
- This suite for VNXe series includes Replicator (iSCSI and NAS).

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.

- This suite for VNX series includes Replication Manager and Data Protection Advisor for Replication Analysis.
- This suite for VNXe series includes Replication Manager.

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 i hosts. EMC Symmetrix VMAX series provides these features and benefits:

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.

IO 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 IOs per second (IOPS). It is very important to understand the IO 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 the following table. Figures are based on an average message size of 75 KB.

Table 1. 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

The URL http://technet.microsoft.com/en-us/library/ee832793.aspx provides information on additional user profiles and their IOPS characteristics.

IO 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 IO has significantly increased.

- For Exchange 2003, the database transactional IO size was 4 KB.
- For Exchange 2007 the database transactional IO size was 8 KB.
- For Exchange Server 2010 the database transactional IO size is 32 KB with a background database maintenance (BDM) read IO size of 256 KB.

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

From a storage perspective, the IO for Exchange Server 2010 is significantly different from previous versions of Exchange. The figure below illustrates the average database transactional IO size observed during Exchange Server 2010 Jetstress testing. Notice that the average read IO is about 120 KB, which is a result of the larger

IO sizes (32 KB database transactional IO size and 256 KB BDM IO size). The same results can be expected for a similar profile in a production environment.

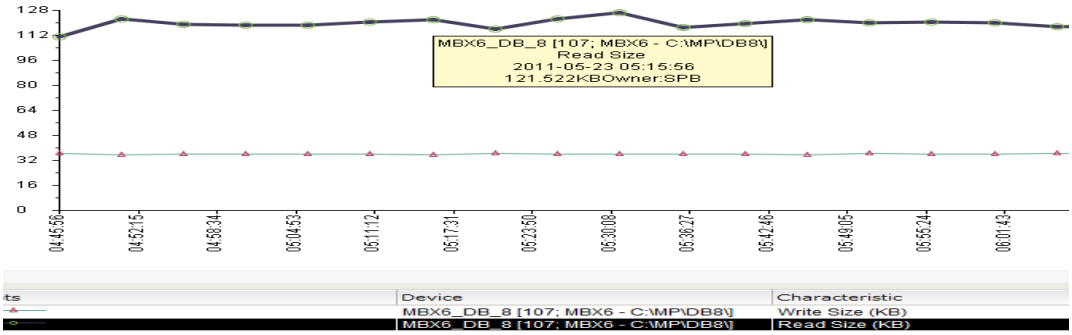

Figure 1. Average database transactional IO size during Exchange Server 2010 Jetstress testing

Background database maintenance

Another important difference in Exchange Server 2010 is the way in which database maintenance is performed, which now includes online defragmentation and online database scanning. The process, called Background Database Maintenance (BDM), produces a large sequential 256 KB read IO. BDM is enabled by default and runs 24x7 on both active and passive database copies.

Important! BDM processes can be scheduled to run against active copies, but cannot be scheduled or disabled for passive DAG copies.

The IO and throughput requirements of BDM needs to be considered as part of storage design when a large number of databases are deployed.

The large 256 KB BDM read IO is factored into EMC's IO calculation in the form of an additional EMC recommended 20 percent overhead factor, which is described in detail in the IO calculations for number of disks required section of this document, which begins on page 24.

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 28, 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 0.15 IOPS users with BDM disabled, which averaged around 3.11 Mb/s.

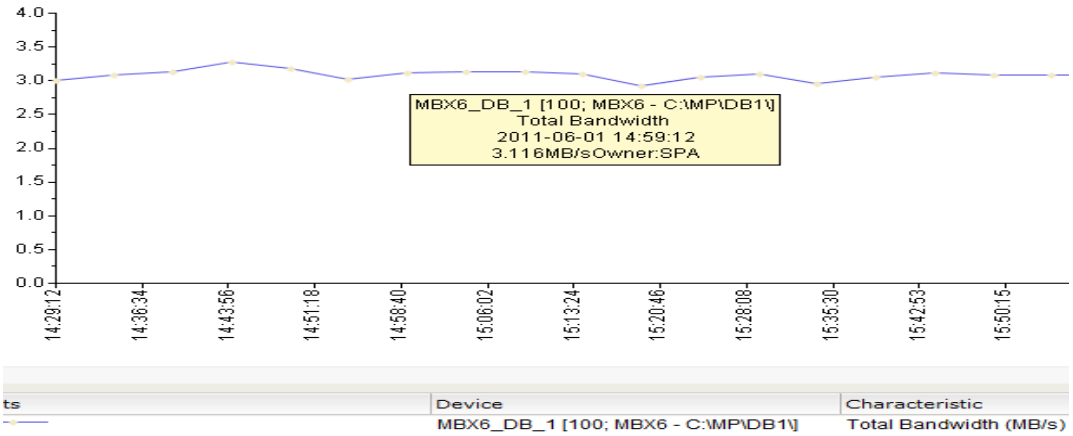

Figure 2. Database throughput for 500 0.15 IOPS users with BDM disabled

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 BDM when performing storage design for Exchange Server 2010.

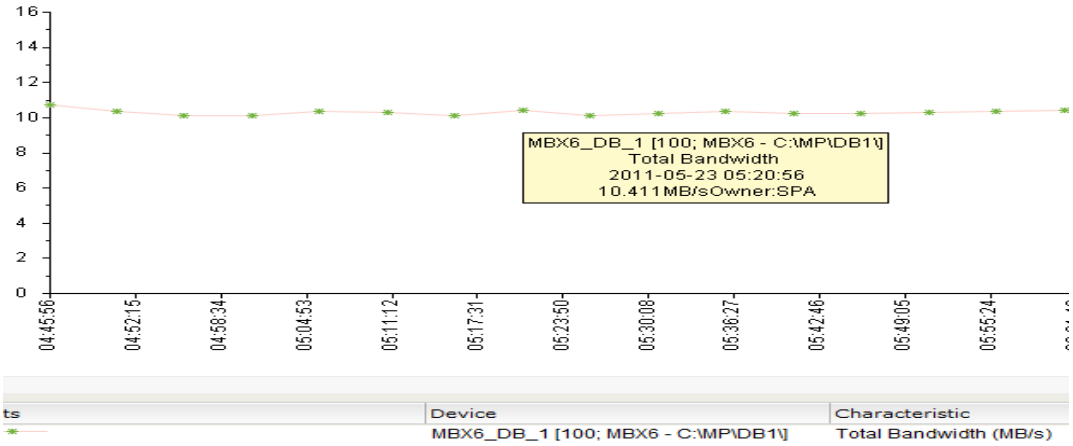

Figure 3. Database throughput for 500 0.15 IOPS users with BDM enabled

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 or 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**—Reliable disk drives with high read/write speeds. These disks are ideal for high IO requirements but might not be appropriate for high capacity requirements.
- **SAS**—An improvement of SCSI. SAS disks provide high capacity with moderate IO speed, which makes them highly suitable for Exchange Server 2010 environments with high IOPS per user requirements.
- **SATA**—The less demanding IO requirements of Exchange Server 2010 make SATA disks a viable option for many Exchange environments. These disks are appropriate for large mailboxes with average IO requirements.
- **NL SAS** – As with SATA disks, NL SAS disks are a good fit for the less demanding IO requirements of Exchange Server 2010. NL SAS 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 IO profiles.
- **Flash**—EMC is a pioneer in Flash drive technology. Flash drives have the highest IO speed with low power consumption. In general, Flash drives are not the best choice for Exchange data, but they can be appropriate when using automated storage tiering features such as EMC FAST VP or FAST Cache to handle any unanticipated IO spikes.

Many IT organizations today are interested in deploying larger mailboxes for their customers. Since the Exchange Server 2010 storage engine now has lower IO 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 IO 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 IO 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.

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 2 when calculating Exchange 2010 IOPS requirements for deployment on VNX and VMAX storage arrays.

Table 2. 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 IO 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 IO 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 3 shows RAID overhead, performance, and storage utilization information for each RAID type.

Note: The RAID overhead value becomes important when performing IO calculations for number of disks required. For details, see the formula for applying RAID overhead on page 24.

Table 3. 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. 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 IO-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 IO requirements before calculating capacity requirements.
- Select appropriate disk types that meet your IO and capacity requirements.

- 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.
- Deploy each DAG copy on its own set of physical disks.
- 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 20).

VNX-specific storage design guidance

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.

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

The use of FAST Cache on VNX systems neither benefits nor detracts from Exchange performance because of the Exchange IO pattern.

- If FAST Cache is used with Exchange Server 2010, segregate database volumes from log volumes by placing each volume type in a separate storage pool, RAID group, or on different LUNs. Enable FAST Cache only for database volumes.
- It is *not* recommended to use FAST Cache for logs.
- Do not use FAST Cache for the purpose of making a storage system pass Jetstress tests.

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

The use of FAST VP with Exchange Server 2010 on VNX is not recommended at this time because Exchange BDM activity affects storage tiering priorities.

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

Either RAID groups or storage pools can be used for Exchange Server 2010 on VNX. When using storage pools, the following guidelines apply:

- Design and grow storage pools by using the appropriate multiplier for best performance (R1/0 4+4, R5 4+1, R6 6+2).
- Use regular (thick) LUN pools only.

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

Thin LUNs are not recommended at this time for Exchange Server 2010 on VNX. Thin LUNs on VNX are not designed for performance workloads. This recommendation might change with future releases of VNX Operating Environment (OE).

What are some of the most important design considerations for Exchange Server 2010 on VNX?

- Set the storage array page size parameter to 16 KB.
- Whether or not you use FAST Cache, segregate database volumes from log volumes by placing each volume type in a separate storage pool, RAID group, or on different LUNs.
- If you use FAST Cache, enable it only for database volumes. Do *not* enable FAST Cache for log volumes.
- When using RAID 5, do not stripe metaLUNs across multiple RAID groups. Doing so can reduce performance, since Exchange Server 2010 is a high-IO and high-bandwidth application.

Symmetrix VMAX-specific storage design guidance

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.

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 IO 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.

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.

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

Whether to use FAST VP with Exchange Server 2010 on VMAX 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 IO issue resolution.

- When using FAST VP, it is *not* recommended to place database files and log files in the same pool, since log files have a low IO requirement and do not need to be moved to a higher tier.
- Keep logs out of the FAST VP policy by placing 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 IO on Fibre Channel disks and 20 percent of IO 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.

What are some of the most important design considerations for Exchange Server 2010 on VMAX?

- When creating LUNs, use fewer but larger hypervolumes 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 metavolumes.
- When using thin LUNs, use concatenated metavolumes.
- Keep logs out of the FAST VP policy by placing log files in a separate storage group from database files.

Virtualization best practices for Exchange Server 2010 on EMC storage

In recent years, virtualizing 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, and flexibility. With the newer, more powerful servers, a six-to-one server consolidation ratio can be achieved, 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. 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 IO 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: 256 SCSI disks per host, processor limits (VMware vSphere 4.1: 8 vCPUs, Windows 2008 Hyper-V: 4 vCPUs), and memory limits.
- Use both hypervisor and OS performance tools to monitor the VMs.

- Use appropriate volume types. For VMware, use VMFS, RDM, or iSCSI on the host or guest. For Hyper-V, use an SCSI pass-through device, VHD, or iSCSI on the host or guest.
- Configurations not currently supported by Microsoft: Hypervisor Snaps, › 2:1 vCPU to pCPU ratio, applications running at the root, VSS backup of root for pass-through disks, or iSCSI disks connected to an initiator on the guest.

The URL http://technet.microsoft.com/en-us/library/aa996719.aspx provides additional information and recommendations for virtualizing Exchange Server 2010 services.

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. 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 (IO, 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.

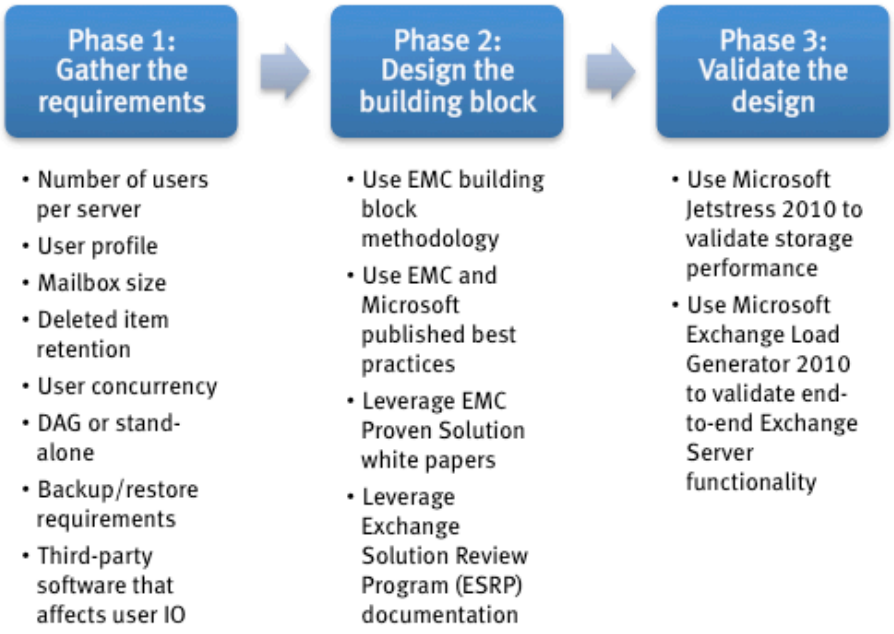

Figure 4. Phases in Exchange building block design

Focus of following sections

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 IO, 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 IO profile (number of messages sent/received per day)
- Mailbox size limit
- Read/write ratio
- Average message size
- Outlook mode
- 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 IO (for example, Blackberry)

Example customer requirements

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

Table 4. 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	Online mode, 100 percent MAPI
Number of mailbox servers per building block	Eight
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	Eight
Number of users per database	500
Deleted items retention (DIR) period	14 days
Log protection buffer (to protect against log truncation failure), in days	Three days
24 x 7 BDM configuration	Enabled
Database read/write ratio	3:2 (60/40 percent) in a mailbox resiliency configuration
User concurrency requirements	100 percent
Third-party software that affects space or IO (for example, Blackberry)	n/a

IO calculations for number of disks required

Phase two 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.

IO calculation

The first step in determining the storage resources required for your building block is to calculate disk requirements based on IO. 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› * ‹user profile› + ‹any additional 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 * ‹user profile› 0.15 + ‹any additional overhead› 0 = 600

Note: Microsoft typically recommends a 20 percent IO 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 = ‹user IOPS› 600 + (‹user IOPS› * 0.20) 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 IO

Next, apply the appropriate RAID overhead value (from Table 3 on page 15) to the write IO, and then divide the result by the appropriate Exchange Server 2010 IOPS per disk value from Table 2 on page 14.

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) = ‹total IOPS› / 65 = ‹total number of disks required per building block›

RAID type affects number of disks required

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 IO and space. The following table presents the equation with various RAID types. Based on this data, RAID 1/0 is the best choice from an IO perspective.

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

In most situations, using the Microsoft Mailbox Server Role Requirements Calculator is the easiest way to calculate space requirements.

- http://msexchangeteam.com/archive/2009/11/09/453117.aspx

While the Microsoft calculator is one option, many people 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 the following URL:

- http://technet.microsoft.com/en-us/library/ee832789.aspx

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 * 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 * 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

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 * 1,728 MB * 1.2 = 1,037 GB

Database LUN size formula

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

- Content index is typically 10 percent of the database size
- Free space requirement is typically 20 percent

Example: 1,037 GB + 103.7 = 1,141 / 0.8 = 1,426 GB per database LUN

Database size	Content index size	LUN free space	Database LUN size
1,037 GB	103.7 GB	285 GB	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)

Here is the formula with sample values and a solution added (angle-bracketed descriptors now shown in grey):

Log LUN size = ‹logs per day, per mailbox› 30 logs at 1 MB per log * ‹number of mailboxes› 500 * ‹protection buffer from log truncation failure, in days› 3 * ‹free space requirement factor› 1.2 = 54 GB

Total space required per building block

Table 5 presents the example database and log LUN size requirements and the resulting total space required per building block.

Table 5. 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›) * ‹database count per server› = (1426 + 54 GB) * 8 = 11840 GB

Space 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.

Total number of disks required

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

Optimal example configuration

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.

Total number of disks required for entire environment

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 = 16 * 8 = 128 disks to support 16,000 users 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 IO throughput validation to ensure that the array buses do not become an IO 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 + ‹BDM throughput per database› 7.5 MB

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 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 IO 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 IO 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 IO 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. Results are reviewed by Microsoft and approved solutions are posted on Microsoft.com. View Microsoft approved reports for EMC submitted solutions at http://www.emc.com/esrp.

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.

Data protection for Exchange Server 2010 on EMC storage

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

DAG feature

In the event of a hardware or software failure, multiple database copies in a database availability group enable high availability with fast failover with no data loss. This eliminates the end-user downtime and resulting lost productivity that's a significant cost of recovering from a past point-in-time backup to disk or tape. DAGs can be extended to multiple sites and can provide resilience against datacenter failures.

Exchange Native Data Protection

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

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 Replication Manager (EMC SnapView™ and EMC Timefinder®)
- EMC Avamar®
- EMC Networker®

Each product has its own strengths. The decision about which one to choose depends on the individual customer's requirements for backup and restore.

EMC hardware-based snap and clone products have been integrated with Microsoft VSS technology for many years. Symmetrix Timefinder and CLARiiON® SnapView enable you to create local point-in-time snapshots and 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 enables the management of EMC point-in-time replication technologies through a centralized management console. 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.

EMC strongly recommends a robust method of enabling rapid Exchange database backup and restore. 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.

Disaster recovery options for Exchange 2010

EMC offers various DR options for Exchange Server 2010. The four most popular options are presented in Table 6. Each option has it advantages and disadvantages. The option that is best for you is determined by your specific DR requirements.

Table 6. 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.
Replication Enabler for Exchange (REE)	EMC RecoverPoint	Sync solution. Integrates with Exchange 2010 third party replication API. Replaces native DAG replication.

Additional backup recommendations

- With medium to high user IO 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.
- With Exchange 2010 in a DAG configuration, a consistency check needs to be done only once a week, since BDM runs constantly and each shipped log is checked before it is applied to passive databases.
- 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 IO 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 IO, 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 Exchange expert.