SAN best practices for deploying
Microsoft Exchange with VMware vSphere

By Ananda Sankaran, Chris Almond, and Puneet Dhawan

Following best practices for Dell™ EqualLogic™ SANs can help organizations design, size, and implement an optimized Microsoft® Exchange Server 2010 deployment based on the VMware® vSphere™ virtualization platform.

Many organizations are looking to move to Microsoft Exchange Server 2010 to help lower total cost of ownership (TCO) while taking advantage of capabilities that enhance performance, reliability, and scalability. To optimize mailbox resiliency features such as database availability groups (DAGs) and online mailbox moves, it is critical to design and size the database storage infrastructure to meet the environment's storage and I/O performance requirements.

Although architectural changes in the Exchange Server 2010 I/O stack make local or direct attach storage (DAS) an interesting option for hosting Exchange data, the inherent challenges of scaling DAS capacity and performance as well as the overhead of managing silos of storage pools cannot be dismissed. The overall cost of DAS may be attractive initially, but over time, the complexity and cost to manage, maintain, and scale DAS quickly increase. According to a July 2009 study by the Enterprise Strategy Group, the TCO over five years of DAS was about 1.4 times

more than the TCO of a comparable Dell EqualLogic Internet SCSI (iSCSI) storage area network (SAN) solution.[1] Using an EqualLogic SAN in a virtualized Exchange environment can provide multiple benefits, including a dynamic infrastructure that supports quick provisioning and flexible deployment; high performance; seamless capacity and performance scalability with a single point of management; comprehensive Exchange-aware data protection and granular data recovery; network path redundancy and I/O load balancing; and a storage platform designed with fully redundant components to help ensure that the messaging infrastructure is available without any compromise.

In July 2010, Dell engineers performed a series of tests to develop storage sizing guidance and demonstrate best practices for designing and running Exchange Server 2010 in a virtualized environment based on VMware vSphere 4.1, Dell PowerEdge™ blade servers, and a Dell EqualLogic PS Series SAN. Solution architects, storage network engineers, system administrators, and IT managers can take advantage of

the practices outlined in this article to help them design, size, and implement an optimized Exchange Server 2010 infrastructure in their own environment.[2]
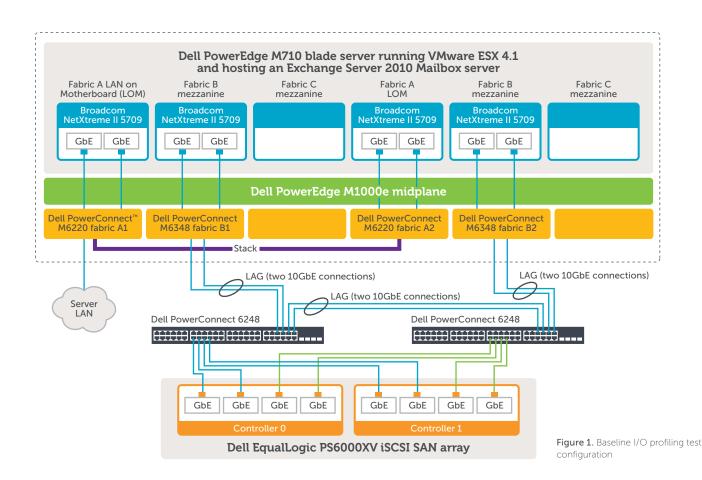
Test environment and results
Figure 1 shows the major system components used in the tests. Following a baseline test to determine the I/O profile of this configuration, the Dell team carried out several rounds of testing designed to evaluate the effect that array type, RAID configuration, database volume layout, and number of arrays had on storage I/O performance. The Microsoft Windows Server® 2008 R2 iSCSI software initiator within the virtual machine (VM) guest OS was used to connect to the external SAN.

Array type
To help meet different organizational needs, the EqualLogic PS Series SAN family includes multiple array models that use a range of different disk types and numbers of disks per array. The results of this test can help solution architects choose the appropriate PS Series array in accordance

Figure 1. Baseline I/O profiling test configuration

with the target Exchange I/Os per second (IOPS). This test measured storage I/O performance for three EqualLogic array types and configurations: an EqualLogic PS6000XV with sixteen 15,000 rpm, 450 GB Serial Attached SCSI (SAS) drives; an EqualLogic PS6500E with forty-eight 7,200 rpm, 1,024 GB Serial ATA (SATA) II drives; and an EqualLogic PS6000X with sixteen 10,000 rpm, 600 GB SAS drives.

As shown in Figure 2, the EqualLogic PS6000XV array provided more IOPS than the EqualLogic PS6000X array, which was due to the 15,000 rpm SAS drives. The EqualLogic PS6500E array with the SATA II drives provided similar IOPS to the PS6000XV, which was due to the increased drive count. Random IOPS performance over time typically depends on the rotational speed of the available drives; in general, faster drives yield more IOPS.

RAID configuration

Although Exchange Server 2010 can run on storage that is not protected using RAID, availability at the storage layer is the key to minimizing downtime and simplifying resolution in failure scenarios. The choice of RAID type and its impact on Exchange performance are important aspects that solution architects must consider.

This test measured storage I/O performance for RAID-10 and RAID-50 volumes on an EqualLogic PS6000XV array and an EqualLogic PS6500E array. Figure 3 shows the results: using RAID-10 on the EqualLogic PS6000XV provided approximately 25 percent more IOPS than using RAID-50, while using RAID-10 on the EqualLogic PS6500E provided approximately 33 percent more IOPS than using RAID-50.

Database volume layout

While designing the layout of Exchange databases and logs, solution architects often face questions on how performance is affected by the placement of Exchange data in storage volumes or logical units (LUNs). This test measured storage I/O performance for different database and log volume placement strategies. The volume layouts were as follows:

- **Layout 1:** Four database volumes (1,250 users, 1,024 GB each) and four log volumes (150 GB each)
- **Layout 2:** Four combined database and log volumes (1,250 users, 1,024 GB each)
- **Layout 3:** Two database volumes (2,500 users, 2,048 GB each) and two log volumes (300 GB each)
- **Layout 4:** One database volume (5,000 users, 4,096 GB each) and one log volume (600 GB)
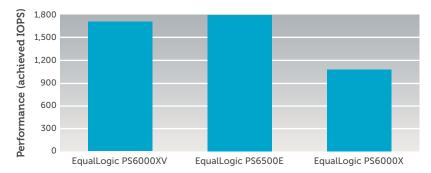
Figure 2. Average storage I/O performance for different array types configured in RAID-50
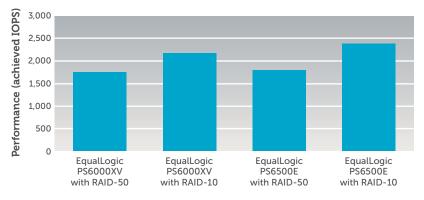

Figure 3. Average storage I/O performance for different RAID configurations
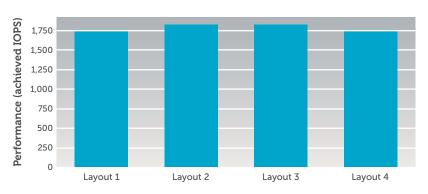

Figure 4. Average storage I/O performance for different RAID-50 database volume layouts
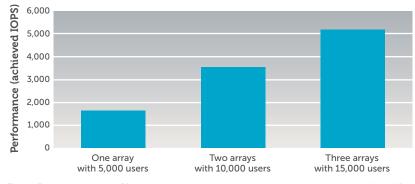

Figure 5. Average storage I/O performance scalability with increases in the number of RAID-50 arrays in a pool

All tests were performed on an EqualLogic PS6000XV array configured in RAID-50. As Figure 4 shows, all four configurations achieved IOPS within a 5 percent range, indicating that for a given number of databases, placement within one or more volumes does not cause a noticeable change in average IOPS.

Number of arrays

Often, the need to support more users and data increases with time. To help avoid increases in complexity and costs, it is imperative to choose a storage architecture that can linearly scale in both capacity and performance without introducing management overhead. Built with this goal in mind, EqualLogic SANs fully meet this requirement.

To measure how effectively Exchange performance scales when adding storage resources, this test measured storage I/O performance as the number of EqualLogic arrays (members) in a pool increases. The tested configurations used one, two, or three EqualLogic PS6000XV members, configured in RAID-50, within the pool. The user mailbox transactional I/O load was increased linearly during the tests along with database size, keeping the database count constant. As Figure 5 shows, the EqualLogic peer storage architecture was able to scale the available storage resources linearly to provide the required IOPS. The number of network interfaces on the server side was kept constant, and the network switches were able to support the arrays under test. If additional arrays are deployed, a different network switch configuration may be required based on port count requirements and workload characteristics.

Recommended best practices

This section provides a summary of best practices developed by Dell engineers to run Microsoft Exchange on top of an infrastructure based on a Dell EqualLogic SAN, Dell PowerEdge servers, and the VMware vSphere virtualization platform. Each of the component layers in the stack requires careful design and configuration to ensure optimal system performance. Within each layer, administrators must consider how the

design affects availability, scalability, and performance.

Best practices for configuring the EqualLogic storage include the following:

- Use an appropriate disk drive speed, number of disk drives, RAID policy, and array model, and be sure to take into account the arrays' usable capacity, Microsoft guidelines for Exchange sizing, database and log capacity, and database and log I/O performance.
- When using a DAG, host active and passive database copies on separate disk drives—for example, by storing the active database and log volumes in one EqualLogic pool, and the passive or copy database and log volumes in a separate EqualLogic pool. If databases are deployed in a DAG, then hosting the database and log volumes on the same physical drives would be acceptable without compromising recoverability.
- Deploy separate volumes for databases and logs. Doing so supports the isolation of respective entities for backup and storage-based recovery options such as snapshots and clones.
- Host the EqualLogic storage volumes as Microsoft Windows® basic disks. Globally Unique Identifier (GUID) Partition Table (GPT) is recommended when configuring partitions for Exchange Server 2010 volumes. Using an NT File System (NTFS) allocation unit size of 64 KB is also a best practice.

Best practices for designing the SAN infrastructure include the following:

- Design redundant SAN component architectures. This design should include the network interface cards (NICs) on the servers and switches for the storage network (including both the server blade chassis switches and the external switches).

- Ensure that the server NIC ports and storage NIC ports are connected such that a single component failure in the SAN does not disable access to any storage array volumes.
- Enable flow control on both the server NICs and the switch ports connecting to server and storage ports.
- Enable jumbo frames on the server ports and the switch ports. Exchange Server 2010 storage I/O is primarily random in nature and uses a 32 KB block size; however, nightly database maintenance and backup operations create sequential I/O patterns. Within VMware ESX, be sure to also set the jumbo frame maximum transmission unit (MTU) value on the virtual switches and the vmkernel interfaces, including for the virtual NIC device used by the VM guest OS.
- Disable Spanning Tree Protocol (STP) on switch ports connecting to end devices like server ports and storage ports, and enable PortFast for these ports.

Best practices for configuring PowerEdge server resources include the following:

- Configure the servers to provide dedicated NICs for Exchange client/ server I/O, and separate dedicated NICs or host bus adapters (HBAs) for iSCSI SAN I/O.
- Provision at least two NICs per server for each path to provide redundancy.
- If possible, segregate Exchange Mailbox server replication traffic onto separate dedicated NICs. This approach helps ensure adequate connection bandwidth for replication operations in a DAG, and helps increase replication performance where database changes caused by user operations are high.
- Provision additional NICs or HBAs for iSCSI SAN I/O as needed based on throughput and IOPS requirements.

Best practices for configuring VMware ESX hosts include the following:

- To take advantage of EqualLogic-aware connection and path management, load balancing, and enhanced performance, install the EqualLogic Host Integration Tools (HIT) for Microsoft Windows in the VM and the EqualLogic Multipathing Extension Module (MEM) plug-in for vSphere on ESX hosts connecting to an EqualLogic SAN.
- Deploy guest operating systems for Exchange server VMs on physical disk drives separate from those hosting the Exchange server data, per Exchange storage requirements.
- Do not use the VMware High Availability (HA), Distributed Resource Scheduler (DRS), and Fault Tolerance (FT) features for Exchange Mailbox server VMs. Host-based availability solutions such as these are not supported by Microsoft for the Mailbox server role.
- Configure separate virtual switches for VM network traffic and iSCSI storage traffic on the ESX hosts. When using both guest and host software initiators, configure separate virtual switches for each connection path.
- Enable jumbo frames on the iSCSI virtual switch. At least two server NICs dedicated to iSCSI traffic must be configured as uplink NICs to the iSCSI virtual switch. When using the ESX iSCSI initiator or iSCSI HBAs on the host, install the EqualLogic MEM plug-in for vSphere.

Best practices for configuring the VMs and guest operating systems include the following:

- Allocate sufficient processor, memory, and disk space resources to the VM and guest OS.
- Apart from using load balancing and automated iSCSI connection management, to take advantage

of Exchange Server 2010–aware EqualLogic protection features such as application-consistent snapshots, clones, and tools such as Auto-Snapshot Manager / Microsoft Edition (ASM/ME), use the iSCSI initiator within the VM to connect to Exchange data and log volumes. Enable the host software initiator along with the EqualLogic MEM plug-in for other applications or VMs when using a guest OS iSCSI initiator.

- When using the guest OS iSCSI initiator, create virtual NICs of type vmxnet3 within the VM; enable jumbo frames, TCP segment offload (TSO), and large receive offload (LRO) in the VM NICs for iSCSI traffic; and use the EqualLogic Multipath I/O (MPIO) Device Specific Module (DSM) installed as part of the EqualLogic HIT in the guest OS.
- Create one virtual NIC for each uplink to the iSCSI virtual switch (for iSCSI traffic from VM to SAN). A given VM cannot use more than one physical uplink on a switch if it has only one virtual NIC.
- In addition to hosting the guest operating systems and data on separate physical drives, configure the guest OS system drives with at least 15 GB of space plus the size of the memory assigned to the guest VM, to reserve disk space for guest memory swaps.

- Install the latest version of VMware Tools in the guest OS.

Best practices for configuring the Exchange servers include the following:

- In general, deploy database copies using a DAG for high availability and resiliency.
- Implement a backup strategy that includes scheduled copying of databases to secondary storage. ASM/ME can support instant Exchange protection and recovery options using application-consistent snapshots and clones; it also integrates with the Windows Volume Shadow Copy Service (VSS) framework to provide application-consistent data copies on the EqualLogic SAN when used with the guest OS iSCSI initiator.

Optimized Exchange infrastructure

Microsoft Exchange Server 2010 and VMware vSphere 4.1 each incorporate features designed to increase the performance, reliability, and scalability of messaging infrastructures. By following the best practices outlined in this article and deploying these platforms on a foundation of Dell EqualLogic PS Series iSCSI SAN arrays and Dell PowerEdge blade servers, organizations can create a robust, reliable, highly scalable Exchange infrastructure to help meet enterprise messaging needs. PS

Ananda Sankaran is a senior development engineer on the Dell Storage Infrastructure and Solutions team, and has more than nine years of experience in software engineering, systems engineering, and solutions architecture.

Chris Almond is a senior consultant responsible for development of publications produced by the Dell Storage Infrastructure and Solutions team.

Puneet Dhawan is a marketing consultant on the Dell Storage Product Marketing team.

Learn more

Dell EqualLogic PS Series:
dell.com/psseries
dell.com/equallogic

Dell and Microsoft Exchange:
dell.com/exchange

Dell and VMware:
dell.com/vmware