

BUSINESS READY CONFIGURATIONS FOR VIRTUALIZATION

FOR DELL™ POWEREDGE™ BLADE
SERVERS, DELL EQUALLOGIC™ STORAGE,
AND VMWARE INFRASTRUCTURE™

May 2009

Dell Virtualization Solutions Engineering
www.dell.com/virtualization



Business-Ready Configurations for Virtualization

Information in this document is subject to change without notice.

© Copyright 2009 Dell Inc. All rights reserved.

Reproduction in any manner whatsoever without the written permission of Dell Inc. is strictly forbidden.

This white paper is for informational purposes only and may contain typographical errors or technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

Dell, the *DELL* logo, *EqualLogic*, *PowerEdge*, *PowerConnect*, and *OpenManage* are trademarks of Dell Inc. *Microsoft* is a registered trademark of Microsoft Corporation; *VMware*, *VMware Infrastructure*, *vCenter*, and *VMotion* are registered trademarks or trademarks (the "Marks") of VMware, Inc. in the United States and/or other jurisdictions. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Dell disclaims proprietary interest in the marks and names of others.

Contents

1	Introduction	5
2	Audience and Scope	5
3	Overview	5
3.1	VMware Infrastructure 3	6
3.2	Dell PowerEdge Blade Server	7
3.3	Dell EqualLogic PS6000 Series iSCSI Arrays	7
3.4	Why Dell blades + VMware + Dell EqualLogic?	8
4	Reference Architecture	9
4.1	Design Principles	9
4.2	Solution Configuration	10
5	Network Architecture	10
5.1	LAN Architecture	11
5.1.1	Traffic Isolation using VLANs	12
5.1.2	Load Balancing and Failover	12
5.1.3	Stacking	13
5.1.4	Uplinks	13
5.2	iSCSI SAN Architecture	14
5.2.1	Storage Connectivity	14
5.2.2	VLAN	15
5.2.3	Load Balancing	15
5.2.4	Stacking	18
5.2.5	Uplinks	18
5.2.6	EqualLogic Network Requirements	19
6	Storage Configuration	19
6.1	Volume Size Considerations	19
6.2	Array RAID Configuration	19
7	Management	20
7.1	Virtualization Management	20
7.2	Systems Management	20
7.3	Storage Management	21
8	How to order the Solution Bundle	21
8.1	Dell Global Services	22
9	Deploying and Configuring the Solution	24
9.1	Rack, Stack, and Cable	24
9.2	Initial Setup	24
10	References	26

Table of Figures

Figure 1: Overview of Dell Blade Server, Dell EqualLogic SAN, and VMware Infrastructure.....6
Figure 2: Reference Architecture.....9
Figure 3: Overview of Network Configuration11
Figure 4: LAN Configuration12
Figure 5: Stacking of I/O modules.....13
Figure 6: iSCSI SAN Configuration.....14
Figure 7: Storage Connectivity15
Figure 8: No load balancing in the case of a single volume16
Figure 9: Multi-pathing with more than one volume.....16
Figure 10: Load balancing of a iSCSI SAN18
Figure 11: Management components.....20

1 Introduction

The Business Ready Configurations for Dell PowerEdge™ blade servers, Dell EqualLogic™ SAN, and VMware Infrastructure™ provides a detailed reference architecture for deploying and using VMware virtualization on Dell blades and iSCSI storage environments. Based on extensive engineering work in architectural design and certification, customers can quickly and confidently deploy these engineering proven architectures into production environments, thereby helping to eliminate much of the costly and time consuming trial and error often encountered during complex infrastructure design and implementation. The solution is optimally configured to run virtualized workloads and is designed to provide redundancy with no single point-of-failure, as well as scalability and manageability. It includes the network architectures, storage configurations, and best practices necessary for deploying and configuring the solution. The solution will help customers achieve the full benefits of VMware Infrastructure, Dell PowerEdge blade servers, and Dell EqualLogic PS Series arrays.

2 Audience and Scope

The intended audience for this white paper is IT administrators, IT managers, and channel partners who are planning to deploy VMware virtualization using Dell blade servers and Dell EqualLogic SAN. This white paper provides an overview of the recommended servers, storage, software, and services. It can be used to plan, scope, and procure the required components to set up a virtualization infrastructure.

This white paper provides reference architecture and best practices for deploying and configuring Dell's 11th generation blade server – M610, Dell EqualLogic PS6000 Series arrays, and VMware Infrastructure 3. The solution uses iSCSI as the storage environment. Fibre Channel is not discussed in this white paper. The solution is designed to be standalone, with the storage device directly connecting to the blade enclosure. Guidelines are provided to connect the storage to existing infrastructure. Based on customer requirements, further customization of the recommended architecture may be required.

3 Overview

This section (Figure 1) provides a high-level product overview of the VMware Infrastructure, Dell PowerEdge blade servers, and Dell EqualLogic PS Series arrays.



VMware Infrastructure 3

- VMware VMotion
- VMware DRS
- VMware HA
- VMware Storage VMotion

Dell PowerEdge Blade Servers

- Energy efficient PowerEdge M1000e enclosure
- 11th generation M610 blade server
 - Intel “Nehalem” processor
 - Intel QuickPath memory technology
 - Unified Server Configurator
- Flex Address
- CMC and iKVM for enclosure management

Dell EqualLogic PS6000 series iSCSI Arrays

- (4) Gigabit Network ports per controller
- 2 GB cache per controller
- Intelligent and Automated management tools
- Self Managing Array
- Seamless scalability
- Top-tier data protection software

Figure 1: Overview of Dell Blade Server, Dell EqualLogic SAN, and VMware Infrastructure

3.1 VMware Infrastructure 3

VMware Infrastructure 3 is comprised of VMware ESX/ESXi and vCenter Server, and includes the following capabilities:

VMware VMotion: VMware VMotion technology provides real-time migration of running virtual machines from one server to another with no disruption or downtime. VMotion can be used to achieve manual load balancing and zero downtime maintenance.

VMware Distributed Resource Scheduler (DRS): VMware DRS technology enables VMotion to automatically achieve load balancing according to resource requirements.

VMware High Availability (HA): VMware HA provides high availability at the virtual machine level. Upon server failure, VMware HA automatically re-starts virtual machines on other physical servers.

VMware Update Manager: VMware Update Manager automates patch management; enforcing compliance to patch standards for VMware ESX hosts, as well as Microsoft or Linux virtual machines. Secure patching of offline virtual machines, as well as automated snapshots and rollback capabilities, reduce risk in the environment. Integration with DRS allows for non-disruptive VMware ESX patching with little-to-no administrative overhead.

VMware Storage VMotion: VMware Storage VMotion enables real-time migration of running virtual machine disks from one storage array to another with no disruption or downtime. It minimizes service disruptions due to planned storage downtime previously incurred for rebalancing or retiring storage arrays.

For more information on VMware Infrastructure, see <http://www.vmware.com/products/vi/>.

3.2 Dell PowerEdge Blade Server

Blade Modular Enclosure: The Dell PowerEdge M1000e is a high-density, energy-efficient blade chassis that supports up to sixteen half-height blade servers, or eight full-height blade servers, and six I/O modules for the blade servers. A high-speed passive mid-plane connects the server modules to the I/O modules, management, and power in the rear of the enclosure. The enclosure includes a flip-out LCD screen for local configuration. The enclosure also includes six hot-pluggable and redundant power supplies and nine hot-pluggable N+1 redundant fan modules.

Blade Servers: The Dell PowerEdge M1000e supports the recently announced 11th generation Dell PowerEdge M610 and M710 blade servers based on the new Intel[®] *Nehalem* processors. The Dell PowerEdge M610 supports 12 DIMM slots and the Dell PowerEdge M710 supports 18 DIMM slots. The blade servers use Intel QuickPath technology to provide a high-speed link to the memory modules. The 11th generation blade servers come with the next generation system management tool: Unified Server Configurator (USC). This helps customers reduce operating costs by simplifying deployment and management. USC supports diagnostics, firmware updates, and hardware configuration. The enclosure also supports the AMD-based M605, M805, and M905 blade servers. The Intel-based M600 legacy blade server is also supported.

I/O Modules: The enclosure provides three redundant fabrics using six I/O modules. The modules can be populated with Cisco Catalyst blade switches, Dell PowerConnect[™] switches, and pass-through modules. Fibre Channel and InfiniBand[®] switch modules are also supported.

Management: The Dell PowerEdge M1000e has integrated management through a redundant Chassis Management Controller (CMC) module for enclosure management and integrated keyboard, video, and mouse (iKVM) modules. Through CMC, the enclosure supports FlexAddress technology. FlexAddress technology allows any M-Series blade enclosure to lock the World Wide Name (WWN) of the Fibre Channel controller and Media Access Control (MAC) of the Ethernet controller into a blade slot. This enables seamless swapping or upgrade of blade servers with Ethernet and Fibre Channel controllers without affecting the LAN/SAN.

For more information on Dell blade servers, see www.dell.com/blades.

3.3 Dell EqualLogic PS6000 Series iSCSI Arrays

The Dell EqualLogic PS6000 is the latest iSCSI storage device from Dell. Its features include four network ports per controller, faster processors, 2 GB cache per controller, support for RAID 6, increased drive capacity, and a new monitoring application, SAN HQ, at no additional cost.

In addition to the new features described above, Dell EqualLogic SAN devices provide the following capabilities:

Reliability: Dell EqualLogic PS6000 Series arrays have hot-swappable redundant components, a choice of RAID types, and hot-spare disks. They also include the Auto-Stat Disk Monitoring System (ADMS) which proactively scans disk drives in the background to help detect media anomalies and correct them.

Scalability: As each array is added to the storage group, the storage capacity and performance, in terms of both bandwidth and IOPS, are increased. This increased capacity can be utilized without downtime. Thin-provisioning permits predefinition of a group of VMFS volumes which are larger than the physical space, allowing for physical space to be added and utilized without downtime when necessary.

Self-Managing Arrays: The arrays offer many self-managing features such as automatic load balancing and storage tiering. A single storage pool can have different models that offer a range of capacity and performance parameters. In addition, different arrays in a storage pool can be configured with different RAID levels, and volumes will automatically be migrated between the RAID levels based on performance data and usage patterns. All data and volume movement can be performed online with zero downtime.

Top-Tier Data Protection Software: Advanced data protection features such as Auto Replication and Auto-Snapshot Manager come standard. The Auto-Snapshot Manager integrates with VMware vCenter and VMware's native snapshot technology to provide intelligent and consistent snapshots.

VMware Site Recovery Manager (SRM) is tightly integrated with the Dell EqualLogic PS Series to offer centralized disaster recovery management, automation, and testing for the entire virtualized datacenter using the PS Series' native Auto-Replication feature integrated directly into Site Recovery Manager using specialized storage adapter software.

For more information on Dell EqualLogic storage, see www.dell.com/equallogic.

3.4 Why Dell blades + VMware + Dell EqualLogic?

VMware Infrastructure on Dell blade servers and Dell EqualLogic SAN provide a compelling infrastructure that meets the demands of today's datacenters. The combined solution offers customers rapid deployment, optimized resource utilization, simplified management, seamless scalability, and energy efficiency with no compromise in performance.

Rapid Deployment

The solution offers several tools that help customers go from an out-of-the-box environment to a production environment quickly and seamlessly. Dell's USC provides a single interface to configure the server BIOS, network controller, and storage controller. USC also supports VMware ESX installations. If customers opt for VMware ESXi, Dell ships servers with integrated ESXi on an embedded SD card. Dell EqualLogic SAN provides a highly simplified deployment process. Dell EqualLogic PS Series arrays go from out-of-the-box to production deployment in under an hour with an intuitive, wizard-based GUI to integrate easily into the existing infrastructure. Using virtualization capabilities, new VMs can be created or existing VMs can be cloned or deployed from templates with just a few clicks.

Optimized Resource Utilization

In addition to increased resource utilization through server consolidation using VMware ESX/ESXi, VMware VMotion and DRS technologies dynamically distribute workloads in all servers in the cluster, ensuring each server's computing resources are optimally used. The Dell EqualLogic SAN complements this by optimally using the storage devices. The virtual machine data is intelligently distributed between all the members (storage devices) in the storage pool, ensuring that all the storage controllers and network interfaces are optimally used. This distribution is done automatically and is completely transparent from the user.

Simplified Management

One of the key benefits of virtualization is simplified management. VMware ESXi software comes integrated with Dell OpenManage™ management software, providing customers with a simplified systems management interface. The Dell EqualLogic PS Series arrays provide a single-pane management view of the entire virtualized storage pool in the SAN. The Dell EqualLogic PS6000 Series arrays intelligently balance workloads across the available arrays with no manual intervention.

Seamless Scalability

VMware virtualization and Dell servers and storage provide a highly scalable solution to meet the growing IT demands of today's datacenters. New server resources can be easily added to VMware DRS clusters. Virtual Machines will automatically migrate to utilize the newly added resources. Similarly, new storage devices can be easily added to Dell EqualLogic storage pools. Dell EqualLogic will automatically migrate the virtual machine data to the newly added storage resource with no downtime.

Energy Efficiency

The Dell PowerEdge M1000e enclosure takes advantage of thermal design efficiencies such as ultra-efficient power supplies and dynamic power-efficient fans with optimized airflow design to efficiently cool the chassis and enable better performance in a lower power envelope.

Consolidating workloads running on older servers on the new energy-efficient Dell blade servers greatly reduces power consumption. This capability, combined with VMware’s DPM technology (currently experimental), which shuts down under-utilized servers and other virtualization benefits, substantially improves the energy efficiency of the solution.

4 Reference Architecture

This section describes the high-level reference architecture for the solution.

4.1 Design Principles

The following principles were used during the design of the architecture.

1. **Optimal hardware configuration for virtualized workloads:** The solution is designed with an optimal hardware configuration to support virtualized workloads. Each blade server is configured with sufficient memory and network adapters required for virtualization.
2. **Redundancy with no single point-of-failure:** Redundancy is incorporated in every aspect of the solution, including networking and storage.
3. **Scalability:** The solution is highly scalable and is based on the architecture of the M1000e chassis and the EqualLogic PS6000 Series arrays. Additional blade servers and storage can be added to increase the compute and storage capacity respectively.
4. **Isolated, redundant, and high-performance network design:** The network architecture is designed to support isolation of various traffic types required in a virtualized environment. It is designed to have no single point-of-failure and have optimal performance through NIC teaming and load balancing.

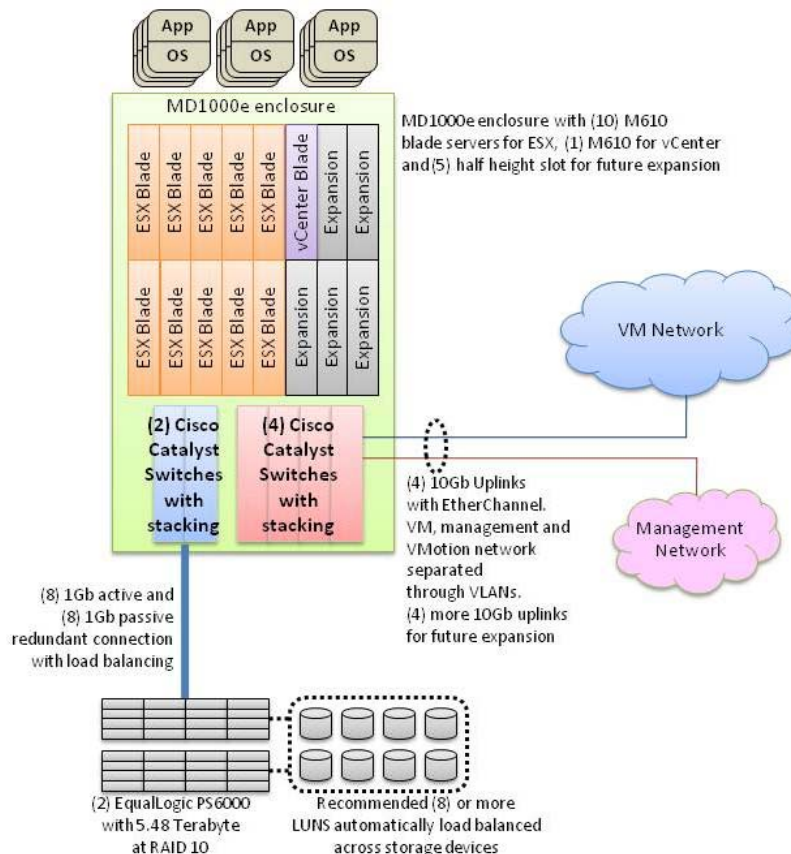


Figure 2: Reference Architecture

4.2 Solution Configuration

The solution consists of a Dell PowerEdge M1000e chassis populated with M610 blade servers and VMware ESX/ESXi. One blade is used to install the VMware vCenter Server. Each blade server is configured as follows:

- Processor: Two Quad-Core Intel Xeon 5500 Series processors (“Nehalem”)
- Memory: 24 to 48 GB
- Network: Two dual-port 1 Gb Ethernet mezzanine cards, in addition to on-board Broadcom dual-port 1Gb Ethernet controller. This provides a total of six network ports.
- Local Storage: Diskless for ESXi and SAS 6iR controller with two drives in mirror configuration.
- Integrated systems management through iDRAC.

The M1000e enclosure has the following configuration:

- Four Cisco Catalyst 3130G/X I/O modules configured for management network and VM network
- Two Cisco Catalyst 3130G I/O modules for the iSCSI network
- Redundant CMC modules for management
- iKVM for remote access

The blade servers are connected to one to two Dell EqualLogic PS6000 Series arrays depending on the storage requirements.

The rest of this document provides guidelines on configuring the solution as illustrated in Figure 2.

5 Network Architecture

The Dell blade chassis has three separate fabrics; namely A, B, and C¹. Each fabric has two I/O modules, making it a total of six I/O modules in the chassis. The I/O modules are A1, A2, B1, B2, C1, and C2. Each I/O module can be an Ethernet physical switch or an Ethernet pass-through module². In this solution, each one of these is an Ethernet fabric containing a Cisco Catalyst 3130G or 3130X blade switch module.

VMware Infrastructure is typically comprised of four types of traffic: virtual machine traffic, management traffic, VMotion traffic, and iSCSI traffic. To support different traffic types in a virtualized environment, create two networks (see Figure 3). The first network is a Local Area Network (LAN), which includes management, virtual machine, and VMotion traffic. It is created by combining I/O modules A1, A2, C1, and C2 using high-speed stacking ports. The second network is dedicated to iSCSI SAN. It is created by combining I/O modules B1 and B2 using high-speed stacking ports³.

Design Principles for Network Architecture: The following design principles are used to develop the network architecture to support the above traffic types:

- **Redundancy:** Both LAN and iSCSI SAN have redundant I/O modules. Redundancy of the network adapters is achieved through NIC teaming at the virtual switch. There is no single point-of-failure.
- **Simplified management through stacking:** Switches servicing the same traffic type is combined into logical fabrics using the high-speed stacking ports on the switches.
- **iSCSI SAN physical isolation:** The iSCSI SAN should be physically separated from the LAN to avoid any network congestion.
- **Logical isolation of VMotion using VLAN:** VMotion traffic is unencrypted. It is important to logically isolate the VMotion traffic using VLANs.
- **Optimal performance:** Load balancing is used to achieve the highest throughput possible.

¹ <http://support.dell.com/support/edocs/systems/pem600/en/hom/HTML/about.htm#wp1212100>

² Fabrics B and C support Fibre Channel and InfiniBand, which are outside the scope of this white paper.

³ There is no special reason for choosing Fabric B for iSCSI network. Fabric C can also be chosen for iSCSI network, with Fabric A and B for LAN.

Based on these principles, we describe how we configure the solution to enable the various traffic types. Best practices for load balancing, failover, and VLAN configuration are provided. This section assumes a certain understanding of networking concepts in a virtualized environment.

See http://www.vmware.com/files/pdf/virtual_networking_concepts.pdf for an overview of VMware's networking capabilities.

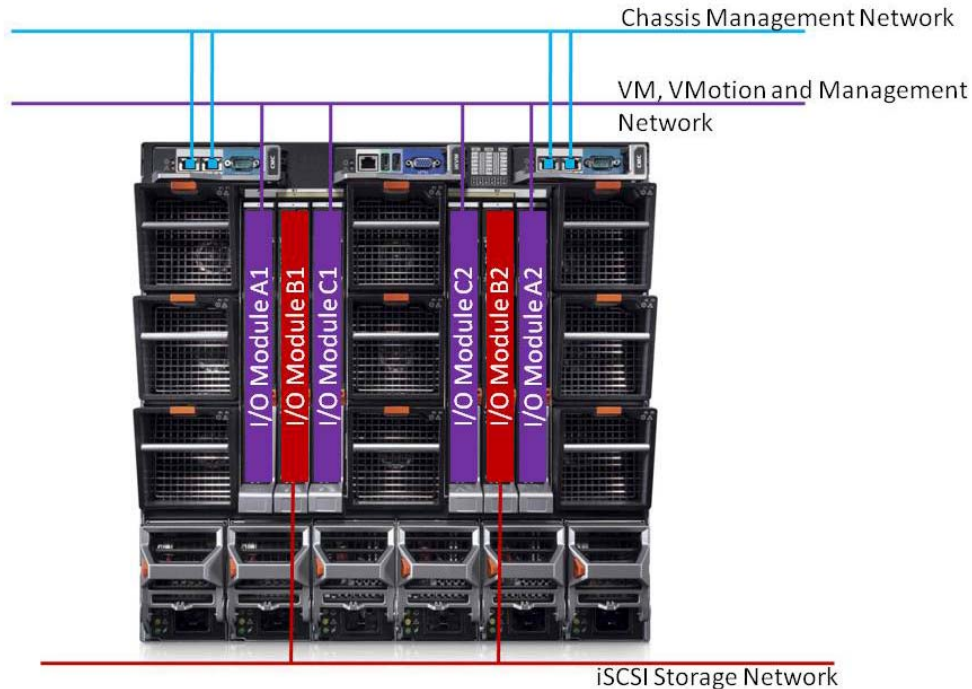


Figure 3: Overview of Network Configuration

5.1 LAN Architecture

The LAN comprises of virtual machine traffic, management traffic, and VMotion traffic. One virtual switch (vSwitch0) is created and four 1 Gb physical network adapters are connected as uplinks to this virtual switch. This creates a team of four network adapters, enabling NIC failover and load balancing. These network adapters are in turn connected to Cisco Catalyst 3130 I/O modules in slots A1, A2, C1, and C2. Three port groups are created on vSwitch0: one for the virtual machine network, another for the VMotion network, and the third for the service console (for ESX) or management (for ESXi). When virtual machines are created, the virtual network adapters are connected to the virtual machine network port group.

Figure 4 illustrates the virtual switch configuration with port groups, and how the virtual switch is connected to physical network adapters and in turn to the I/O modules.

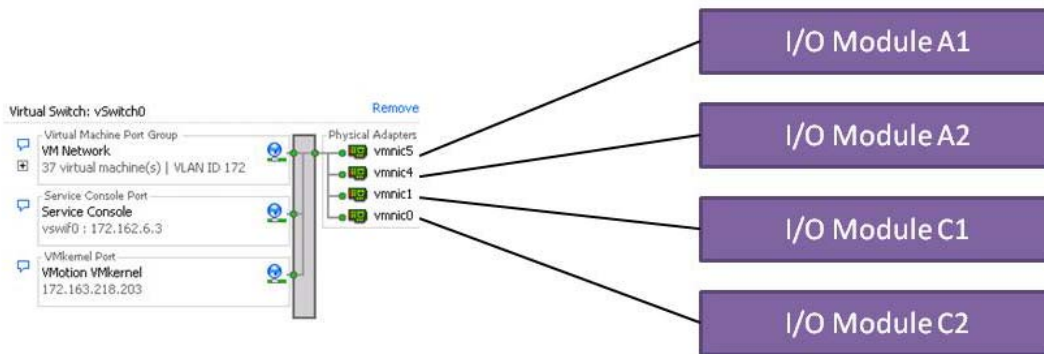


Figure 4: LAN Configuration

5.1.1 Traffic Isolation using VLANs

The traffic on the LAN is separated into three VLANs; one VLAN each for management, VMotion, and virtual machine traffic. Network traffic is tagged with the respective VLAN ID for each traffic type in the virtual switch. This is achieved through the Virtual Switch Tagging (VST) mode. In this mode, a VLAN is assigned to each of the three port groups. The virtual switch port group tags all outbound frames and removes tags for all inbound frames. Routing between VLANs is dependent on the specific customer requirements and is not included in the reference architecture. If desired, the Cisco Catalyst 3130 I/O modules can be easily configured to provide the routing.

Trunking must be used so that all the VLANs can share the same physical connection. To achieve this, all the internal ports in the Cisco I/O modules are configured to be in the trunk mode.

VMotion traffic is unencrypted, and so it is highly recommended to isolate VMotion traffic⁴. Using the above VLAN configuration, we achieve traffic isolation between various traffic types, including the VMotion traffic. The four network adapters provide sufficient bandwidth for all the traffic types.

5.1.2 Load Balancing and Failover

Load balancing enables sharing of network traffic between the network adapters in a team, thereby generating higher throughput. The VMware virtual switch has three options to configure load balancing:

- **Route based on the originating virtual switch port ID (default configuration):** For outbound traffic, the physical adapter is selected based on the hash of the virtual port. This means that a given virtual network adapter will use only one physical adapter at any given time to transmit network packets. Packets are received on the same physical adapter.
- **Route based on source MAC hash:** Here, the physical adapter is selected for transmit based on the hash on the source MAC address. This means that a given virtual network adapter will use only one physical adapter at any given time to transmit network packets. Packets are received on the same physical adapter.
- **Route based on IP hash:** Here, the physical adapter is selected for transmit based on the hash of the source and destination IP address. EtherChannel can be used with this setting. The physical switch load balancing algorithm will then determine which physical adapter is selected to receive the packets. To achieve proper load balancing, both the virtual switches and the physical switches need to be configured. Note that the virtual switch and the physical switch hashing algorithms work independent of each other.

The default load balancing configuration of route based on the originating virtual switch port ID can be used. Multiple VMs will use different physical adapters to transmit and receive traffic. Service console (for ESX) or management network (for ESXi) has only one port ID or MAC address. Hence, the service console or management

⁴ <http://blogs.vmware.com/security/2008/02/keeping-your-vm.html>

network will use only one physical adapter for communicating, unless there is a failover to another physical adapter. The same is the case for VMotion traffic. VMotion will use only one physical adapter for communicating, unless there is a failover to another physical adapter.

Note: All the physical network adapters will be configured as *active* by default when load balancing is enabled.

5.1.3 Stacking

Each switch I/O module is a standalone switch that can be managed independently. Switches can also be stacked together to form a single virtual blade switch (VBS) using high-speed stacking cables. One of the switch modules acts as the master and the entire VBS can be managed by connecting to the master switch. Stacking provides several advantages. By stacking switch modules together, we manage the switches as a single entity; thereby simplifying management. Stacking also provides for a high-speed data path between the switch modules. The Cisco Catalyst switch modules support stacking and up to eight Cisco switches can be stacked to form a single VBS.

Switch modules A1, A2, C1, and C2 can be stacked together to form a VBS. Four stacking cables are used to form a *ring* as shown in Figure 5. This provides redundancy in case of a single-stack link failure. Stacking for the iSCSI SAN is described in the next subsection.

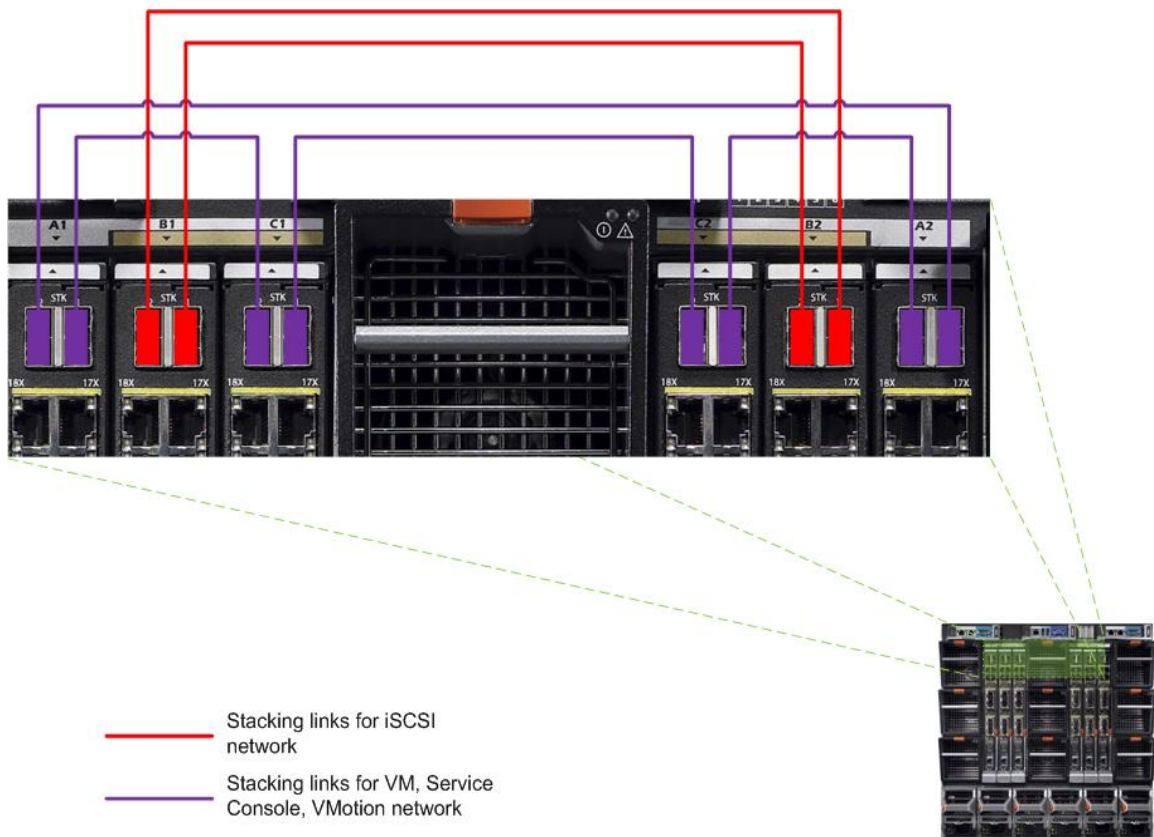


Figure 5: Stacking of I/O modules

5.1.4 Uplinks

The blade chassis can be connected to the existing network infrastructure in the datacenter with various options. Choosing the right option depends on the configuration of the existing infrastructure. In this section, we will discuss the options available and our recommendations for the solution.

Each Cisco 3130X I/O module has four built-in 1 Gbports. Each I/O module can also be populated with two 10 Gb Ethernet uplink ports. This gives us a total of 24 Gb uplink throughput per I/O module. With four I/O modules, this gives us a total of 96 Gb uplink capability.

If the existing infrastructure supports a 10 Gb network, one of the two 10 Gb ports in each Cisco I/O module can be populated. The 10 Gb links from the four switches are joined together into a single EtherChannel to provide a high-speed data path into the blade chassis. The additional 10 Gb port and four 1 Gb ports for each I/O module can be used for future expansion.

If the existing infrastructure does not support 10 Gb, the solution can be configured with Cisco 3130G I/O modules with 1 Gb uplinks instead of Cisco 3130X I/O modules.

VLAN and routing must be configured on these uplinks depending on the existing infrastructure requirements.

Note: Jumbo frames can be used in the LAN. Jumbo frames must be configured end-to-end in the network to fully utilize the capabilities. In VMware environments, jumbo frames must be configured on the virtual switch in addition to configuring jumbo frames on the virtual network adapter or VMkernel interface. For more information on configuring jumbo frames, see the *ESX Server 3 Configuration Guide* at <http://www.vmware.com/support/pubs/>.

5.2 iSCSI SAN Architecture

The iSCSI SAN comprises of iSCSI traffic to the Dell EqualLogic PS Series arrays. In this section, we discuss the best practices for configuring the iSCSI SAN using the software iSCSI initiator in ESX/ESXi. Software initiator in the guest OSES is also supported, but is outside the scope of this document.

A virtual switch (vSwitch1) is created and the two physical ports connected to Fabric B are added. These network adapters are in turn connected to Cisco Catalyst 3130G switches in I/O modules – B1 and B2.

Figure 6 illustrates the virtual switch configuration with port groups and how the virtual switch is connected to the physical network adapters, and in turn to the I/O modules. The figure is based on ESX. If ESXi is used, there is no need to configure the service console port group for iSCSI.

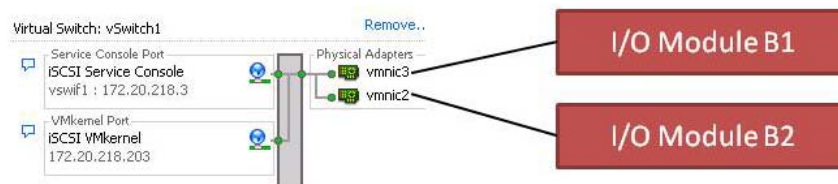


Figure 6: iSCSI SAN Configuration

5.2.1 Storage Connectivity

Each Cisco Catalyst 3130G switch contains four fixed 1 Gb ports. In addition to this, each switch has two TwinGig Converter modules placed in the X2 ports and populated with a total of four 1 Gb SFP modules. This gives us a total of sixteen 1 Gb ports for connecting the storage arrays. Each controller on the Dell EqualLogic PS6000 Series array has eight ports (four active and four passive connections). The four active ports are split between two different physical switches to provide redundancy. The passive ports are connected in a similar fashion.

Figure 7 illustrates how two Dell EqualLogic PS6000 Series arrays are connected to the Dell PowerEdge M1000e chassis.

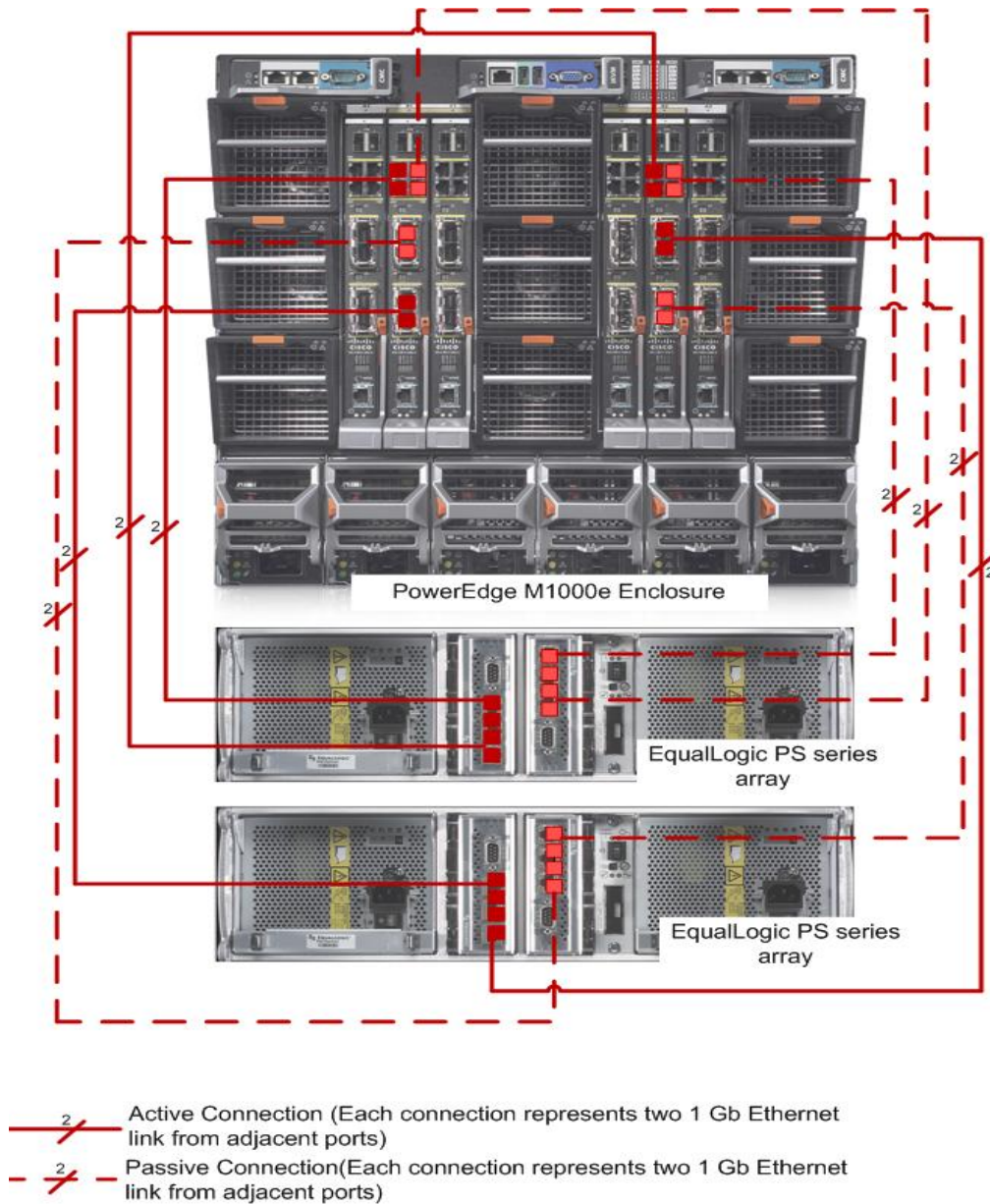


Figure 7: Storage Connectivity

5.2.2 VLAN

The storage traffic is completely isolated from other networks. There is no requirement to enable multiple VLANs for this architecture. The ports on the physical switch can be configured in the access mode. The virtual switch does not have VLANs defined in the port groups.

5.2.3 Load Balancing

To understand load balancing, it is important to understand the basics of iSCSI and how it is implemented in VMware ESX/ESXi. See References for further information.

Here are some key fundamentals that will help you in understanding how load balancing works:

- ESX software's iSCSI initiator establishes only one connection to each target.
- The current ESX version uses only one path for each connection.
- Dell EqualLogic SAN exposes each volume as a separate target with LUN ID 0.

This means that ESX can establish and use only one path to a single volume on Dell EqualLogic SAN, even if multiple paths are available. In other words, for a given volume, only one path is active at any given time. This is illustrated in Figure 8. In this case, redundancy is still available. The path will failover in case of a network adapter or switch failure.

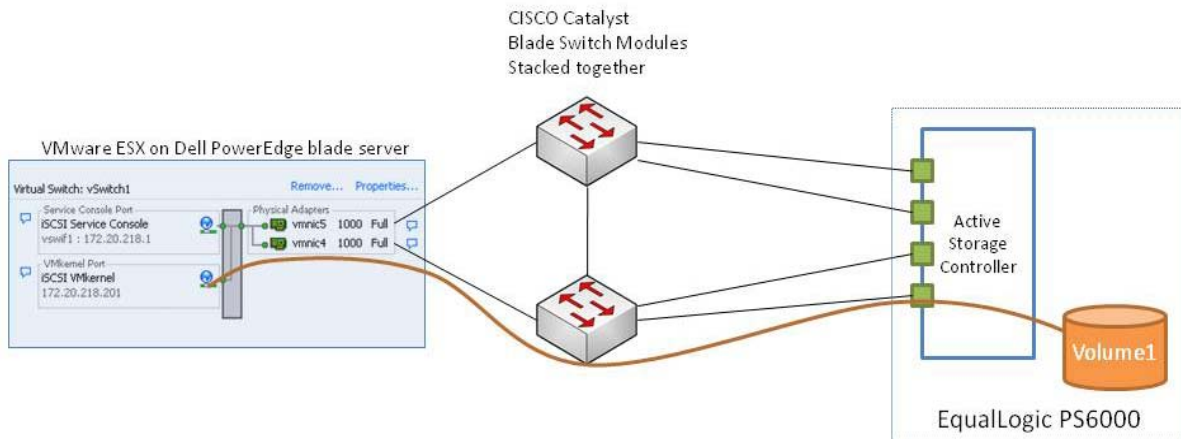


Figure 8: No load balancing in the case of a single volume

If there are multiple volumes, then multiple paths can be used as illustrated in Figure 9.

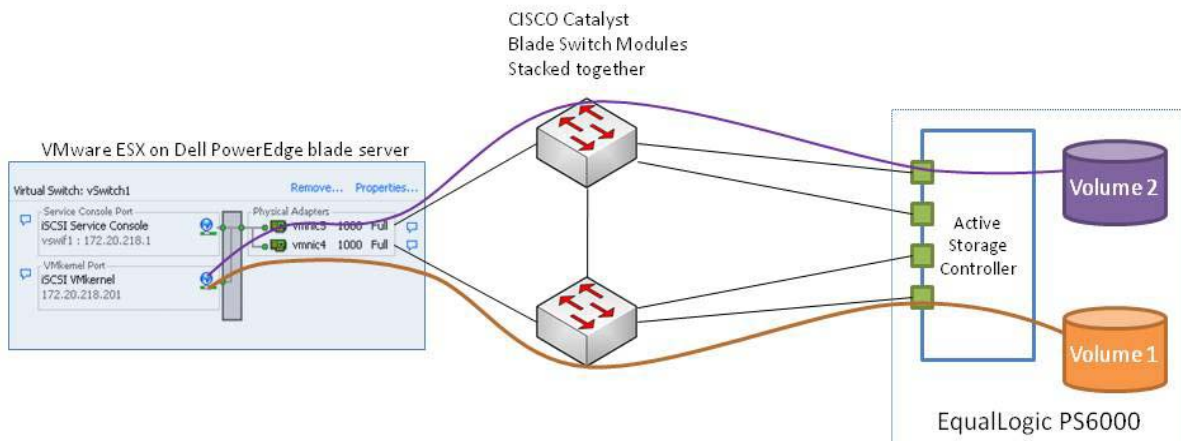


Figure 9: Multi-pathing with more than one volume

In order to achieve this load balancing, both the virtual and Cisco switch modules must be configured.

The VMware virtual switch has three options to configure load balancing:

- **Route based on IP hash - Recommended:** Here, the physical adapter is selected for transmit based on the hash on the source and destination IP address. Software iSCSI uses only one IP address to connect to the Dell EqualLogic SAN. The Dell EqualLogic SAN distributes iSCSI connections to the targets across four active network interfaces. Each interface has a unique IP address. Hence, the hash outcome of source and destination IP address can be different for each volume. If multiple volumes are configured, this could result in using both the physical adapters in each blade server. EtherChannel must be used in the physical switch modules with this setting, and the source destination IP (*src-dst-ip*) hash algorithm must be selected for load balancing. The physical switch load balancing algorithm will then determine which physical adapter is selected to receive the packets.
- **Route based on the originating virtual switch port ID (default configuration) – Not Recommended:** Here, the physical adapter is selected for transmit based on the hash of the virtual port. Software iSCSI uses only one virtual port and selecting this option would mean that only one physical adapter will be used. Hence, this option does not provide load balancing for an iSCSI SAN.
- **Route based on source MAC hash – Not Recommended:** Here, the physical adapter is selected for transmit based on the hash on the source MAC address. Software iSCSI uses only one MAC address and selecting this option would mean that only one physical adapter will be used. Hence, this option does not provide load balancing for an iSCSI SAN.

Dell EqualLogic SAN automatically distributes iSCSI connections between the available four network interfaces in each controller. It also automatically distributes volumes between different storage devices in the storage pool. No additional configuration is required.

By configuring the virtual and Cisco blade switch modules, good load balancing can be achieved⁵. This is illustrated in Figure 10.

⁵ Example switch configuration available in Appendix A - Blade Configuration Scripts

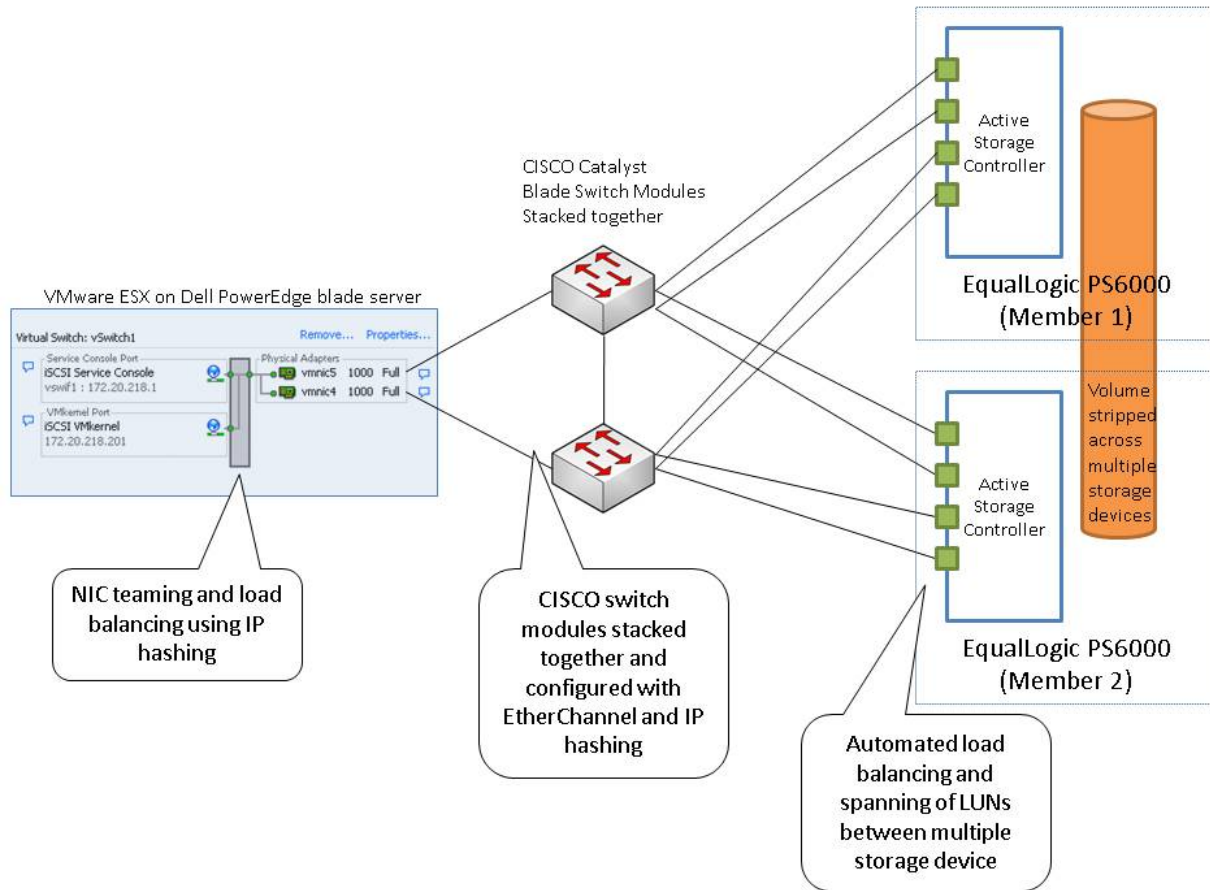


Figure 10: Load balancing of a iSCSI SAN

Note: The distribution of traffic between the available network adapters (or paths) is dependent on the hashing algorithms. The hashing algorithms are more efficient with more elements to hash.

5.2.4 Stacking

For the iSCSI SAN, switch modules B1 and B2 can be stacked together to form a VBS. Two stacking cables are used to form a *ring* as shown in Figure 5. This provides redundancy in case of a single-stack link failure. Stacking also acts as inter switch link (ISL). The stacking cables provide sufficient bandwidth to meet the ISL best practices for the storage arrays.

5.2.5 Uplinks

The reference architecture connects the blade chassis to two EqualLogic storage arrays directly. This is designed for a standalone configuration and there are no available ports to connect the iSCSI network to an existing infrastructure. The EqualLogic storage should be managed from the management blade in the chassis.

If uplinks are required, or the storage needs to be integrated with an external management infrastructure, intermediate switches need to be used. The storage arrays can be connected with a two-tiered architecture. Best practices for using two-tiered storage architecture is outside the scope of this document.

5.2.6 EqualLogic Network Requirements

In addition to the above guidelines, EqualLogic has specific recommendations for connecting PS Series arrays to the network. We have highlighted some of the important recommendations. For more information, see the *Dell EqualLogic PS Quick Start Guide* at <https://www.equallogic.com/support/> (Account registration may be required).

- Do not use Spanning-Tree (STP) on switch ports that connect end nodes (iSCSI initiators or array network interfaces). However, if you want to use STP or Rapid STP (preferable to STP), you should enable the port settings available on some switches that let the port immediately transition into the STP forwarding state upon link up. This functionality can reduce network interruptions that occur when devices restart and should only be enabled on switch ports that connect end nodes.

Note: The use of Spanning-Tree for a single-cable connection between switches is encouraged, as is the use of trunking for multi-cable connections between switches.

- Enable Flow Control on each switch port and NIC that handles iSCSI traffic. PS Series arrays will correctly respond to Flow Control.
- Disable unicast storm control on each switch that handles iSCSI traffic if the switch provides this feature. However, the use of broadcast and multicast storm control is encouraged on switches.

Note: Jumbo frames are not supported in ESX 3.5. Update to ESX 4 for iSCSI software initiator.

6 Storage Configuration

The section provides recommendations for configuring the Storage Array. For storage networking and stacking guidelines, see [Stacking](#).

6.1 Volume Size Considerations

Volume sizes would depend on the customer environment and the type of workloads. It is highly recommended to create multiple volumes to overcome the load balancing behavior described in the previous section. To utilize all the allocated storage ports, at least one volume per active port should be created for a total of eight volumes in the recommended configuration. This would allow ESX to establish multiple connections with Dell EqualLogic SAN and would ensure an even load across the physical ports. Depending on the environment, you may decide to create multiple ~500 GB volumes with multiple VMs. Ensure to include space for Guest OS memory cache, snapshots, and VMware configuration files when sizing the volumes. Additionally, you can configure thin-provisioned volumes to grow on demand only when additional storage is needed for those volumes. Thin-provisioning can increase the efficiency of storage utilization.

6.2 Array RAID Configuration

The storage array RAID configuration is highly dependent on the workload in your virtual environment. The Dell EqualLogic PS Series arrays support four RAID types – RAID 5, RAID 6, RAID 10, and RAID 50. It is recommended to configure the array in RAID 10 for best throughput and performance. In general, RAID 10 provides the best performance, at the expense of storage capacity. RAID 50 generally provides more usable storage, but has less performance than RAID 10 in random I/O situations and requires additional overhead in the case of a drive failure scenario. RAID 5 provides the most storage capacity at the expense of a little lower performance and availability.

7 Management

In this section, we describe the various components used to manage the solution.

Figure 11 displays the various management components.

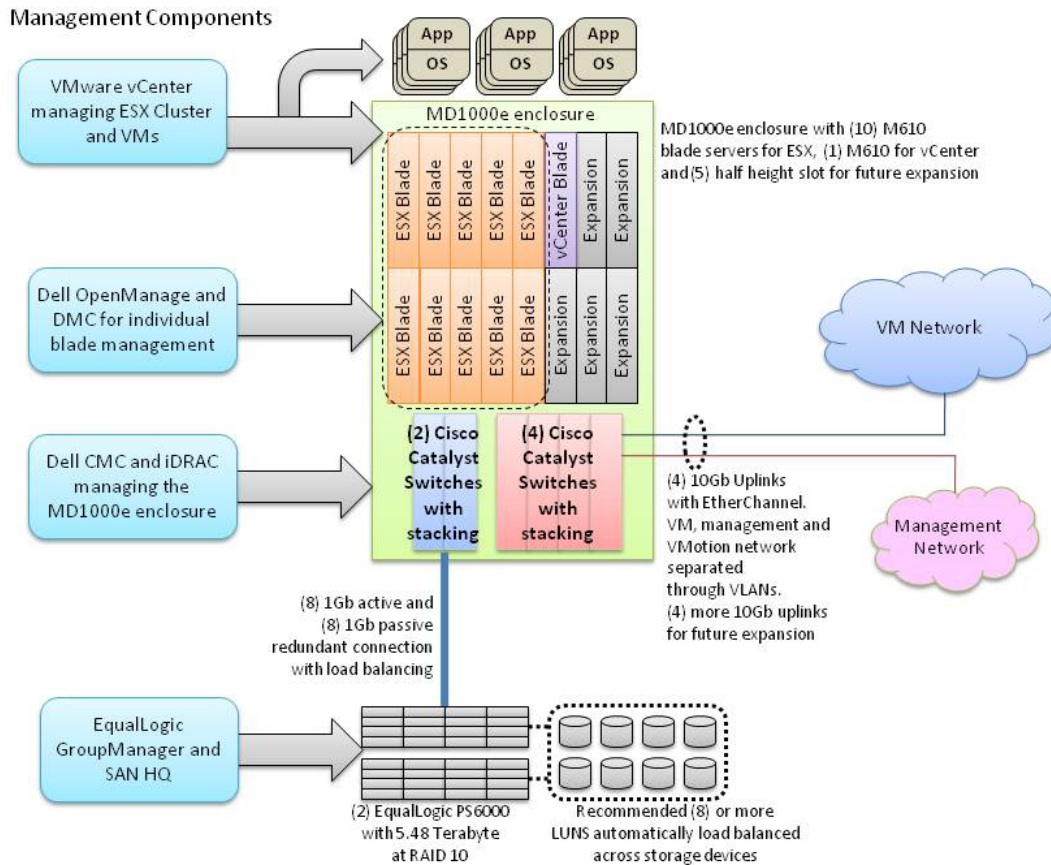


Figure 11: Management components

7.1 Virtualization Management

The VMware vCenter provides a single control and management center for VMware ESX Servers. VMware vCenter allows you to manage all ESX Servers and virtual machines, utilize advanced features such as VMotion, HA, and DRS, monitor performance and resource optimization, and use standard templates to provision new servers in minutes. For more information on vCenter, see <http://www.vmware.com/products/vcenter/>.

PlateSpin[®] Workload Management allows you to profile, migrate, protect, and manage server workloads between physical and virtual infrastructures in heterogeneous IT environments. The PlateSpin Workload Management portfolio for Dell servers gives Dell customers a powerful data center management platform for Dell's 32- and 64-bit Windows and Linux servers with support for leading hypervisors. For more information on PlateSpin migration tools and other products, see www.dell.com/platespin.

7.2 Systems Management

The following tools and software can be used to manage the hardware:

Dell OpenManage: Dell OpenManage Server Administrator (OMSA) can be installed on VMware ESX and can be used to manage the hardware. VMware ESXi comes integrated with Dell OpenManage Server Administrator. For more information on OpenManage and its capabilities, see www.dell.com/openmanage.

Deployment and change management using Unified Server Configurator (USC): The Dell PowerEdge M610 blade servers come with USC. This helps customers reduce operating costs by simplifying deployment and management. Key features include: Diagnostics, self-update (UEFI, Driver Pack update), firmware updates (BIOS, NIC FW, RAID Controllers), and hardware configuration.

One-to-many using DMC: Dell Management Console is Dell’s next generation systems management software. It provides features such as configurable dashboards, graphical reports, role-based security, hierarchical software deployment, MIB Import for management of events from non-Dell devices, one-to-many hardware configuration wizards, and greater scalability through remote database infrastructure.

Out-of-band CMC and iDRAC: CMC provides a single, secure interface to manage the inventory, configuration, monitoring, and alerting for chassis components (iKVM, CMC), I/O modules, servers, and iDRAC. It also provides excellent real-time power management, monitoring, and alerting capabilities. The Dell chassis also provides the users with system-level power limiting, slot-based prioritization, and dynamic power engagement functionalities. iDRAC on each server provides the flexibility to remotely manage the server through Console redirection and Virtual CD-ROM/DVD/Floppy/Flash capabilities.

7.3 Storage Management

Dell EqualLogic provides a rich set of management features that is available at no additional cost and come with exceptionally easy-to-use management tools. For more information on Dell EqualLogic features and management capabilities, see www.dell.com/equallogic.

8 How to order the Solution Bundle

To order the solution discussed in this paper, contact your Dell Sales Representative with the *Solution ID* provided below. Customers can order any of three pre-configured solutions. Customization can be made to the configurations to fit specific needs. Based on the customization, some of the best practices discussed in this white paper may not apply.

Table 1 Solution Specification

	Entry Configuration	Medium Configuration	Large Configuration
Hardware			
Solution ID	689996	690002	690006
Blade Chassis	(1) M1000e	(1) M1000e	(1) M1000e
ESX Blade Server	(4) x M610	(6) x M610	(10) x M610
Storage Device	(1) PS6000X	(2) PS6000XV	(2) PS6000XV
Management Blade Server	(1) x M610	(1) x M610	(1) x M610
Chassis Configuration			
I/O module for A1	Cisco M 3130G	Cisco M 3130G	Cisco M 3130X with (2) X2-CX4 10Gb Module
I/O module for B1	Cisco M 3130G with (2) 1000BAS-T SFP modules	Cisco M 3130G with (4) 1000BAS-T SFP modules	Cisco M 3130G with (4) 1000BAS-T SFP modules
I/O module for C1	Cisco M 3130G	Cisco M 3130G	Cisco M 3130X with (2) X2-CX4 10Gb Module
I/O module for A2	Cisco M 3130G	Cisco M 3130G	Cisco M 3130X with (2) X2-CX4 10Gb Module
I/O module for B2	Cisco M 3130G with (2) 1000BAS-T SFP modules	Cisco M 3130G with (4) 1000BAS-T SFP modules	Cisco M 3130G with (4) 1000BAS-T SFP modules
I/O module for C2	Cisco M 3130G	Cisco M 3130G	Cisco M 3130X with (2) X2-CX4 10Gb Module
Stacking cables	6	6	6

Management	Redundant Chassis Management Controllers (CMC)	Redundant Chassis Management Controllers (CMC)	Redundant Chassis Management Controllers (CMC)
KVM	Integrated Avocent keyboard, video, and mouse (iKVM) switch	Integrated Avocent keyboard, video, and mouse (iKVM) switch	Integrated Avocent keyboard, video, and mouse (iKVM) switch
ESX Blade Server Configuration			
Blade Server Model	M610	M610	M610
Processor	(2) x Intel Xeon (Nehalem) E5520, 2.26Ghz, 8M Cache	(2) x Intel Xeon (Nehalem) E5520, 2.26Ghz, 8M Cache	(2) x Intel Xeon (Nehalem) E5520, 2.26Ghz, 8M Cache
Memory	24 GB (6 x 4 GB, DDR3)	48 GB (12 x 4 GB, DDR3)	48 GB (12 x 4 GB, DDR3)
Add-in Controllers	(2) Broadcom 5709 Dual Port GbE I/O Card for M-Series Blades	(2) Broadcom 5709 Dual Port GbE I/O Card for M-Series Blades	(2) Broadcom 5709 Dual Port GbE I/O Card for M-Series Blades
Local storage and controller	Diskless configuration	Diskless configuration	Diskless configuration
Hypervisor	VMware ESXi v3.5 Enterprise 2S, 3 Year Subscription	VMware ESXi v3.5 Enterprise 2S, 3 Year Subscription	VMware ESXi v3.5 Enterprise 2S, 3 Year Subscription
Storage Configuration			
Storage Device	(1) PS6000X	(2) PS6000XV	(2) PS6000XV
Drives	16 X 400GB, 10K SAS	16 X 450GB, 15K SAS	16 X 450GB, 15K SAS
Storage Capacity	6.4 Terabyte capacity	7.2 Terabyte per device	7.2 Terabyte per device
Management Blade Server			
Blade Server Model	(1) M610	(1) M610	(1) M610
Processor	(2) x Intel Xeon (Nehalem) E5520, 2.26Ghz, 8M Cache	(2) x Intel Xeon (Nehalem) E5520, 2.26Ghz, 8M Cache	(2) x Intel Xeon (Nehalem) E5520, 2.26Ghz, 8M Cache
Memory	8 GB	8 GB	8 GB
Network Controller	(2) Broadcom 5709 Dual Port GbE I/O Card for M-Series Blades	(2) Broadcom 5709 Dual Port GbE I/O Card for M-Series Blades	(2) Broadcom 5709 Dual Port GbE I/O Card for M-Series Blades
Virtualization Management Software	VMware VirtualCenter v2.5, 3Yr Upgrade Subscription	VMware VirtualCenter v2.5, 3Yr Upgrade Subscription	VMware VirtualCenter v2.5, 3Yr Upgrade Subscription
Software and Services			
Additional Software	PlateSpin Migrate for Dell (P2V 5-Pack), License & 1 Year Maintenance	PlateSpin Migrate for Dell (P2V 5-Pack), License & 1 Year Maintenance	PlateSpin Migrate for Dell (P2V 5-Pack), License & 1 Year Maintenance
Services	<ul style="list-style-type: none"> 3 Year ProSupport for IT and Mission Critical 4HR 7x24 Onsite Pack EqualLogics PS Array Installation M1000e Blade Chassis On-site installation service 	<ul style="list-style-type: none"> 3 Year ProSupport for IT and Mission Critical 4HR 7x24 Onsite Pack EqualLogics PS Array Installation M1000e Blade Chassis On-site installation service 	<ul style="list-style-type: none"> 3 Year ProSupport for IT and Mission Critical 4HR 7x24 Onsite Pack EqualLogics PS Array Installation M1000e Blade Chassis On-site installation service

8.1 Dell Global Services

Dell Global Services helps customers find the “better path” to virtualization to reduce total cost of ownership, speed time to ROI, increase agility, and reclaim IT resources. Today’s financial environment dictates that IT reduce its

Business-Ready Configurations for Virtualization

budget while still providing ever increasing infrastructure services. Dell Global Services believes that by heavily leveraging server virtualization, and increasing the rate of adoption of this technology, you can accomplish this task.

Dell Global Services virtualization consultants work with customers to design and plan around the most common bottlenecks in virtualization implementations. Our methodology, tools, and processes are designed to speed up the implementation, ease migration scheduling, automate reporting, and provide transparency to the entire process. Dell Global Services also has the ability to offer end-to-end solutions with a single point-of-contact for hardware, software, services, and on-going support. In a time when many providers aspire to do everything, we at Dell Global Services are focused on IT infrastructure services excellence.

To engage Dell Global Services, see

http://www.dell.com/content/topics/global.aspx/services/adi/virtualization_new?c=us&cs=555&l=en&s=biz.

9 Deploying and Configuring the Solution

This section provides guidelines for the initial deployment and configuration of the solution.

9.1 Rack, Stack, and Cable

1. Rack the chassis with blade servers and I/O modules.
2. Rack the Dell EqualLogic storage.
3. Connect power to the chassis and storage.
4. Consider PDUs. This document provides detailed information on how to configure the PDUs: http://www.dell.com/downloads/global/products/pedge/en/pe_m1000e_selection_whitepaper.pdf
5. Cable servers and storage.
 - a. Cable the server and storage for Ethernet connectivity as per the architecture.
 - b. Connect the switches with the given stacking configuration. Cable all the Ethernet switches with the stacking cables before you power on the chassis for the first time.

9.2 Initial Setup

Configuring the Chassis

Configure the Chassis Management Controller using the web interface or the front LCD panel on the chassis.

- Use static chassis management IP address (preferable).
- Enable FlexAddressing.
- Configure the iDRAC IP address for the blades.
- Configure the Cisco Switch Management IP.
- PowerThreshold.

Follow the procedures in the CMC User Guide

<http://support.dell.com/support/edocs/software/smdrac3/cmc/cmc2.0/en/ug/pdf/cmcug.pdf>) and iDrac User Guide

(<http://support.dell.com/support/edocs/software/smdrac3/idrac/idrac20modular/en/ug/pdf/ug.pdf>)

Configuring Networking

Configure network switch modules. It is recommended that you create a configuration file for the network switches. These files are written in a sparse format using a range of interfaces.

Two ways of loading the switch file onto the switches are recommended:

Serial: You can connect a serial cable from a Microsoft Windows® or Linux computer to the master switch of a VBS, which can be identified by the blue light at the bottom of the switch. You will need a terminal program such as HyperTerminal.

TFTP: You can use the Chassis CMC to configure an IP address for each VBS, then use telnet or ssh to connect to the switch and then tftp to copy the config file to the switch. A TFTP server is required for this method.

Configuring the Server

Configure the server either locally or by using the iDRAC Console Redirection remotely. Perform the following:

- Enable VT for BIOS.
- For ESX, set up RAID 1 with the two drives. For ESXi, it is recommended to have diskless servers.
- Configure Unified Server Configurator for self-update and firmware repository.
- Install VMware ESX 3.5 on the server. (See the *VMware Solution Guide* for further information.)
- Install Dell OpenManage using the System Management media.

See the Unified Server Guide (http://www.dell.com/downloads/global/solutions/unified_server_overview.pdf) for more information.

Configuring Storage

Configure the Dell EqualLogic PS Series. For information on configuring storage, see the *Dell EqualLogic PS QuickStart Guide* at <https://www.equallogic.com/support/> (Account registration may be required).

Configuring vCenter

You will need to install license server and the SQL Database before installing VMware vCenter. See the *ESX QuickStart Guide* at http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_quickstart.pdf to install vCenter on the management server.

Deployment Guidelines

1. Create a cluster in vCenter with DRS and DPM features enabled. Add the ESX servers to the cluster.
2. vCenter Networking (**vCenter > Configuration > Networking**):
 - vSwitch1** (Management, VM Traffic, and VMotion)
 - a. Add the NIC ports connected to Switch A1, A2, C1, and C2.
 - b. Enable load balancing.
 - c. Create port groups for Management, VM Traffic, and VMotion (vmKernel). It is highly recommended to use different VLANs for each port group. For example, Service Console VLAN 162, VM Traffic VLAN 163, and VMotion VLAN 172.
 - vSwitch2** (iSCSI)
 - a. Create vSwitch2 for iSCSI.
 - b. Add the NIC ports connected to Switch B1 and B2.
 - c. Add the Service Console port.
 - d. Add a VMkernel port for iSCSI.
 - e. Enable IP Hash Load Balancing (**vSwitch > Properties > Edit > Properties**).
3. Cisco configuration:
 - a. Stack the switches A1, A2, C1, and C2 in a ring topology (See Figure 5).
 - b. Create EtherChannel groups between the ports connected to each Virtual Switch (in the above step) for each server.
 - c. Ensure that the switch is configured to handle the VLANs (preferably Trunk mode) created for the Service Console, VM Traffic, and VMotion.
4. Dell EqualLogic PS6000XV– Configuring iSCSI:
 - a. Create volumes.
 - b. Create access list:
 - i. Ensure that you add iSCSI Service Console and VMkernel iSCSI IP to the access list. (Use * to denote the IP address range.)
 - ii. Create volumes to be shared between multiple servers (cluster).
 - c. Enable iSCSI in the VMkernel NIC and perform a re-scan to find the volumes.
5. vCenter – Add the volumes as volumes to all the ESX Servers in the cluster.
6. Power Monitoring and Management:
 - a. Use CMC to perform power monitoring of Dell blades when workload VMs are running.
 - b. Use DMC to set up SNMP alerts for physical machines for easy management.

10 References

iSCSI overview - A “Multivendor Post” to help our mutual iSCSI customers using VMware

http://virtualgeek.typepad.com/virtual_geek/2009/01/a-multivendor-post-to-help-our-mutual-iscsi-customers-using-vmware.html

Integrating Blade Solutions with EqualLogic SANs

http://www.dell.com/downloads/global/partnerdirect/apj/Integrating_Blades_to_EqualLogic_SAN.pdf

Cisco Products

http://www.cisco.com/en/US/products/ps6746/Products_Sub_Category_Home.html

Cisco 3130 Product Page

<http://www.cisco.com/en/US/products/ps8764/index.html>

VMware Infrastructure 3 in a Cisco Network Environment

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/vmware/VMware.html

Cisco Catalyst 3750 and 2970 Switches: Using Switches with a PS Series Group

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5269>

Common issues when using the ESX SW iSCSI Initiator

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5271>

Deploying Pools and Tiered Storage in a PS Series SAN

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5239>

Deploying Thin Provisioning in a PS Series SAN

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5245>

Microsoft Windows: Aligning Disk Sectors for Optimal Performance

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5223>

Network Connection and Performance Guidelines

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5311>

Network Connection Guidelines

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5331>

Protecting the Virtual Environment Using Auto-Snapshot Manager/VMware Edition

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=7369>

PS Series Array Network Performance Guidelines

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5229>

Snapshots in a VMware Virtual Infrastructure

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5259>

VMware Virtual Infrastructure 3.x Considerations, Configuration, and Operation Using an EqualLogic PS Series SAN

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5235>

Windows Disk Sector Alignment

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5323>