VMware® vSphere™ Reference Architecture for Small Medium Business

Dell Virtualization Reference Architecture

Dell Virtualization Solutions Engineering

Feedback: solutionfeedback@dell.com

November 2009

Table of Contents

1 Introduction

Reference Architectures for Virtualization are pre-configured, easy to order bundled solutions designed to aid in ordering, deployment and maintenance of a virtual infrastructure. Reference Architectures for Small and Medium Business (SMB) include Dell™ PowerEdge™ Rack servers, Dell EqualLogic™ SAN and VMware® vSphere™. This white paper provides detailed reference architectures and best practices for deploying and configuring a Reference Architecture targeted at SMB. Based on extensive engineering work in architectural design and certification, customers can quickly and confidently deploy these engineering proven architectures into production environments, thereby eliminating much of the 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. This document covers the network architectures, storage configurations, and best practices necessary for deploying and configuring the solution.

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 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.

The solution discussed in this white paper is based on Dell PowerEdge R510, Dell EqualLogic PS4000 Series iSCSI arrays, and VMware vSphere 4 Essentials Plus and Advanced Edition. The solution is focused on SMB deployments. Based on customer requirements, further customization of the recommended architecture might be required.

3 Overview

Figure 1 shows a typical architecture of VMware vSphere on Dell PowerEdge server and Dell EqualLogic storage. In this section we provide an overview of the various components used in such architecture.
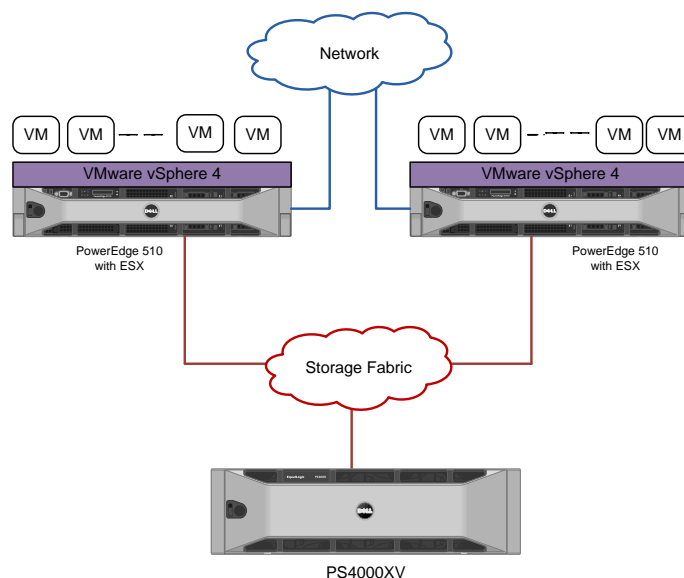

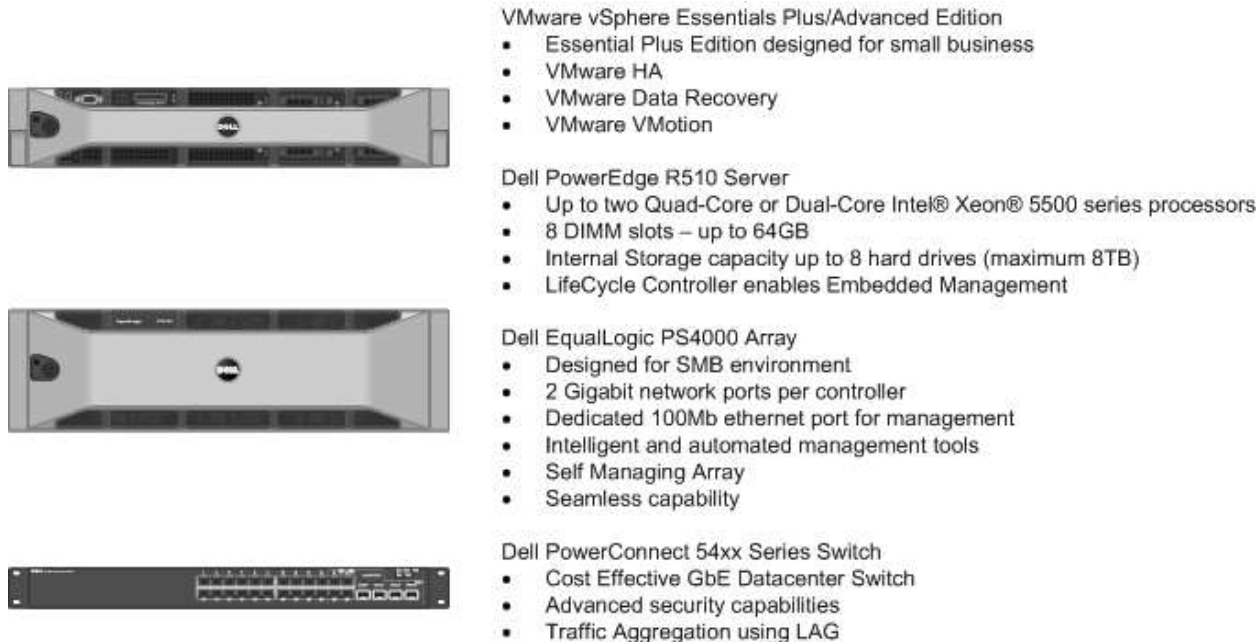**Figure 1: Typical vSphere 4 Architecture with PS Series SAN**

VMware vSphere Essentials Plus/Advanced Edition
- Essential Plus Edition designed for small business
- VMware HA
- VMware Data Recovery
- VMware VMotion

Dell PowerEdge R510 Server
- Up to two Quad-Core or Dual-Core Intel® Xeon® 5500 series processors
- 8 DIMM slots – up to 64GB
- Internal Storage capacity up to 8 hard drives (maximum 8TB)
- LifeCycle Controller enables Embedded Management

Dell EqualLogic PS4000 Array
- Designed for SMB environment
- 2 Gigabit network ports per controller
- Dedicated 100Mb ethernet port for management
- Intelligent and automated management tools
- Self Managing Array
- Seamless capability

Dell PowerConnect 54xx Series Switch
- Cost Effective GbE Datacenter Switch
- Advanced security capabilities
- Traffic Aggregation using LAG

Figure 2: Solution Components

3.1 VMware vSphere 4

VMware vSphere 4 is the latest server virtualization software from VMware. vSphere Essentials Plus and vSphere Advanced Edition are different editions of vSphere suitable for the small and medium business environment. Among the various features in vSphere, features like VMware HA, VMware Data Recovery and VMotion™ are of specific interest while designing and deploying a high availability business ready configuration for small and medium business.

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 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 VMotion is only available with the vSphere Advanced Edition. VMware VMotion is not available in vSphere Essentials Plus Edition.

VMware Data Recovery: VMware data recovery is a disk-based backup and recovery solution providing data protection for the virtual machines. It is fully integrated with VMware vCenter Server which enables centralized management of backup jobs.

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

3.2 Dell PowerEdge R510 Server

The PowerEdge R510 follows the 11th Generation PowerEdge portfolio. R510 is a two socket 2U rack server designed with energy-optimized options supporting Intel® Xeon® 5500 series processors, DDR3 memory and advanced embedded management capability. R510 is available in two different chassis options – four hard drive chassis or eight hard drive chassis. With the eight hard drive chassis option, PowerEdge R510 provides up to 8TB of internal storage. R510 server has one dual-port Broadcom 5716 Gigabit network adapter. The server provides high availability capabilities with hot plug drives, redundant power supplies and redundant cooling.

The Lifecycle Controller is the engine for advanced embedded management and is delivered as part of the optional iDRAC Express or iDRAC Enterprise in the PowerEdge R510. The Lifecycle Controller helps to simplify administrator tasks by performing a complete set of provisioning functions such as system

deployment, system updates, hardware configuration and diagnostics from a single intuitive interface called Unified Server Configurator (USC) in a pre-OS environment.

3.3 Dell EqualLogic PS4000 Series iSCSI SAN

The Dell EqualLogic PS4000 is the latest iSCSI storage device from Dell designed for deployment in a small to medium business environment. It consists of two 1Gb Ethernet network ports per controller for data and one 10/100 Mb port dedicated for management network, 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 PS4000 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™ Server and VMware's native snapshot technology to provide intelligent and consistent snapshots.

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

3.4 Dell PowerConnect 5400 Series Switches

Dell PowerConnect 5400 series switches deliver 24 ports (PowerConnect 5424) or 48 ports (PowerConnect 5448) of wire-speed Gigabit Ethernet with advanced security and enterprise management features to help meet the needs of organizations of all sizes. To provide availability at the network layer, redundant Ethernet switches can be used in combination with NIC teaming on the ESX hosts to provide protection against the failure of a switch or a network device. PowerConnect 5400 series supports VLANs and up to eight Gigabit ports can be combined into a LAG, providing an aggregated bandwidth of 8 Gbps.

4 Specification

In the following table we provide the specification of three pre-configured solutions in this Reference Architecture for SMB; Stand-alone, High Availability (HA), and High Availability + VMotion. In the rest of this document, we will discuss reference architectures and best practices for deploying the solution.

Table 1: Solution Specification

	Stand-alone	High Availability	High Availability + VMotion
Configuration Summary			
Virtualization Server	(1) x R510	(2) x R510	(3) x R510
Virtualization Software¹	VMware vSphere 4 – Essentials Plus, which includes: • High Availability • Data Recovery • VMware vCenter Server	VMware vSphere 4 – Essentials Plus, which includes: • High Availability • Data Recovery • VMware vCenter Server	VMware vSphere 4 – Advanced, which includes: • High Availability • Data Recovery • VMware vCenter Server • VMware VMotion • VMware Fault Tolerance • vShield Zones • Hot Add
Storage Device	Local Storage (8 Internal HDD)	(1) x PS4000	(1) x PS4000
Management Server	None	(1) x R410	(1) x R410
Network Configuration	None	(2) x PowerConnect 5424 Network Switches	(4) x PowerConnect 5424 Network Switches
Virtualization Server Configuration			
Server Model	PowerEdge R510	PowerEdge R510	PowerEdge R510
Processor	(2) x Intel Xeon E5520, 2.26Ghz, 8M Cache	(2) x Intel Xeon E5520, 2.26Ghz, 8M Cache	(2) x Intel Xeon E5520, 2.26Ghz, 8M Cache
Memory	32 GB (8 x 4GB, DDR3)	32 GB (8 x 4 GB, DDR3)	32 GB (8 x 4 GB, DDR3)

---

[1] Only features relevant to the solution are discussed in this white paper.

Add-in Network Controllers	(2) x Broadcom 5709 Dual Port GbE PCI/e Card	(2) x Broadcom 5709 Dual Port GbE PCI/e Card	(2) x Broadcom 5709 Dual Port GbE PCI/e card
Internal Storage Controller	PERC 6/I	PERC 6/I	PERC 6/I
Hard disk	(8) x 300GB, 15K SAS (RAID 5)	(2) x 146GB, 15K SAS (RAID1)	(2) x 146GB, 15K SAS (RAID1)
Hypervisor per server	VMware vSphere 4.0 Essentials Plus (ESX)	VMware vSphere 4.0 Essentials Plus (ESX)	VMware vSphere 4.0 Advanced (ESX)
Embedded Management	iDRAC 6 Express	iDRAC 6 Express	iDRAC 6 Enterprise
Server Management	OpenManage 6.1.2	OpenManage 6.1.2	OpenManage 6.1.2
External Storage Configuration			
Storage Device	None	(1) x PS4000	(1) x PS4000
Drives	N/A	(16) x 450GB, 15K SAS	(16) x 450GB, 15K SAS
Raw Storage Capacity	N/A	7.2 Terabyte per array	7.2 Terabyte per array
Management Server[2]			
Rack Server Model	None	(1) x R410	(1) x R410
Processor	N/A	(1) x Intel Xeon E5520, 2.26Ghz, 8M Cache	(1) x Intel Xeon E5520, 2.26Ghz, 8M Cache
Memory	N/A	8 GB	8 GB
Internal Storage	N/A	SAS 6/iR; (2) x 146GB, 15K SAS Drives	SAS 6/iR; (2) x 146GB, 15K SAS Drives
Operating System	N/A	Windows Server 2008 x64 Standard Edition	Windows Server 2008 x64 Standard Edition

---

[2] VMware vCenter Server can also be installed in a Virtual Machine. For more information, refer: www.vmware.com/pdf/vi3_vc_in_vm.pdf

Virtualization Management Software	N/A	VMware vCenter Server for Essentials (as a part of the VMware vSphere Essentials license)	VMware vCenter Server 4.0, 3Yr Upgrade Subscription
Embedded Management	N/A	iDRAC Express	iDRAC Enterprise
Additional Software and Services			
Services	• 3 Year ProSupport for IT and Mission Critical 4HR 7x24 Onsite Pack	• 3 Year ProSupport for IT and Mission Critical 4HR 7x24 Onsite Pack • EqualLogic PS Array Installation	• 3 Year ProSupport for IT and Mission Critical 4HR 7x24 Onsite Pack • EqualLogic PS Array Installation
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

Important note about licensing: Each VMware vSphere Essentials Plus licensing comes includes:

- ESX Licenses for 3 servers (with 1 or 2 sockets per server)
- One vCenter Server for Essentials License

For Stand-alone Configuration, we have included VMware vSphere Essentials Plus as a part of R510 server configuration. This enables customers to install ESX Essentials on two additional servers (with 1 or 2 sockets) and install VMware vCenter for Essentials with no additional licensing cost.

For High Availability Configuration, we have included VMware vSphere Essentials Plus as a part of R410 server configuration. Customers can install ESX Essentials on the two R510 servers and install vCenter Server on the R410 server. Customers can also install ESX Essentials on one additional server (with 1 or 2 sockets) with no additional licensing cost.

High Availability + VMotion Configuration includes R510s, each with VMware vSphere Advanced Edition. The solution also includes R410 with VMware vCenter Server Edition.

5 Reference Architecture

This section will describe the reference architecture for the solution. As mentioned in the previous section, we provide three pre-configured solution for this Reference Architecture: Stand-alone, High Availability and High Availability with VMotion.

5.1 Design Principles

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

1. **Optimal Configuration for SMB:** The solution components for this Reference Architecture are specifically designed for the requirements of SMB environment. VMware vSphere software Edition and the hardware components are selected to aid in virtualization for SMB.
2. **Optimal Server Configuration for Virtualization:** The solution is designed with an optimal hardware configuration to support virtualized workloads. Each server is configured with sufficient memory and network adapters required for virtualization.
3. **Redundancy with no single point-of-failure:** Redundancy is incorporated in every aspect of the solution, including networking and storage to avoid any single of failure.
4. **Isolated and Redundant Network Architecture:** 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.

In the rest of the section, we provide an overview of the reference architecture for the three configurations.

5.2 Stand-alone Configuration

This configuration can be used for providing an entry level virtualization solution for a small business environment. The configuration consists of single PowerEdge R510 with VMware vSphere Essentials Plus licensing. In this configuration the internal storage is used for storing the virtual machine files. R510 supports up to a maximum of 8TB of internal storage with the 8 drive chassis option. For management, VMware vCenter Server can be installed in a PE R410 or installed in a virtual machine.

5.3 High Availability Configuration

Figure 3 shows the architecture diagram for High Availability configuration for small business environment. The configuration consists of two PowerEdge R510 servers running vSphere Essentials Plus connected to one EqualLogic PS4000 array in an iSCSI SAN environment. The network infrastructure consists of two PowerConnect 5424 switches. PowerEdge R410 server is used for VMware vCenter Server.

More details for the network and storage architecture will be provided in the following sections.
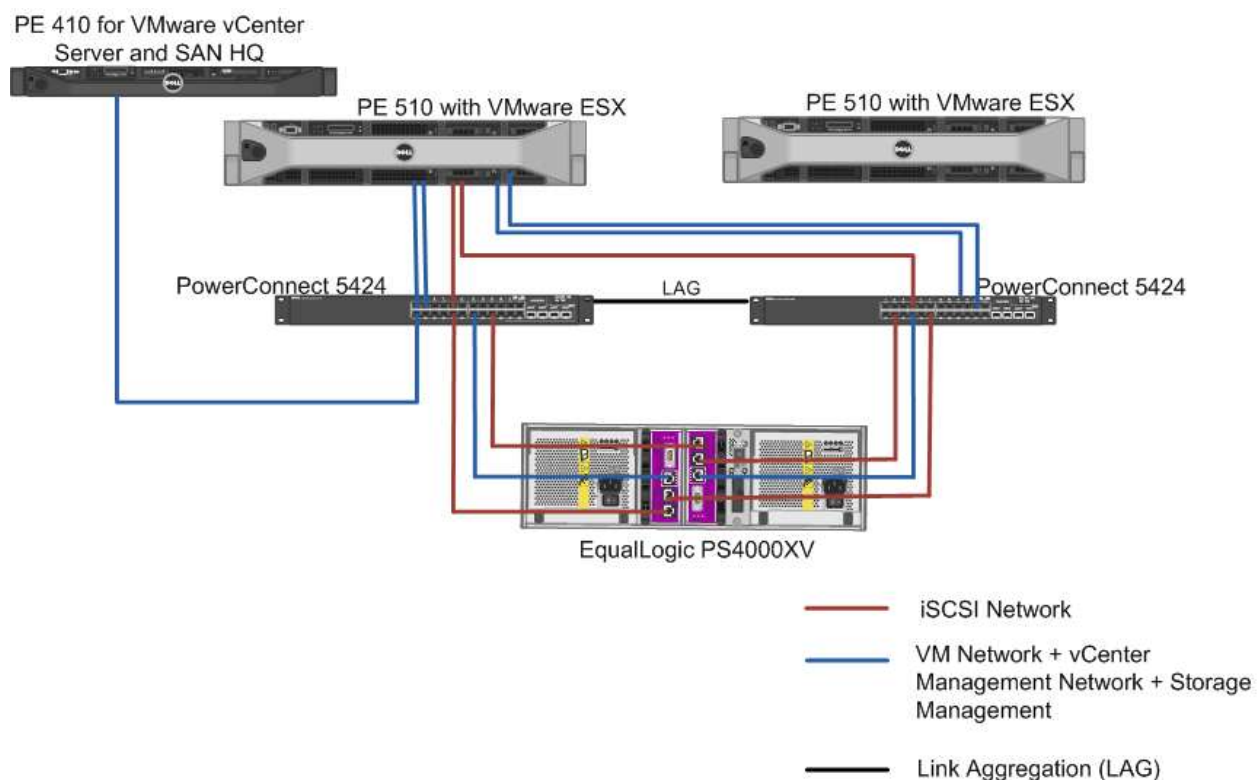
Figure 3: Reference Architecture for High Availability Configuration

5.4 High Availability + VMotion Configuration

This configuration is based on VMware vSphere Advanced Edition. Figure 4 shows the architectural diagram. The solution consists of three Dell PowerEdge R510 servers running VMware ESX (vSphere Advanced Edition) connected to one PS4000 series array in an iSCSI SAN environment. Two PowerConnect 5424 switches are used to support Virtual Machine network, Management network and VMkernel network traffic. Two more PowerConnect 5424 switches are dedicated for iSCSI traffic. We have dedicated separate switches for iSCSI network to allow future addition of new iSCSI storage devices, enabling scalability. PowerEdge R410 server is used for deploying vCenter Server.

More details for the network and storage architecture will be provided in the following sections.
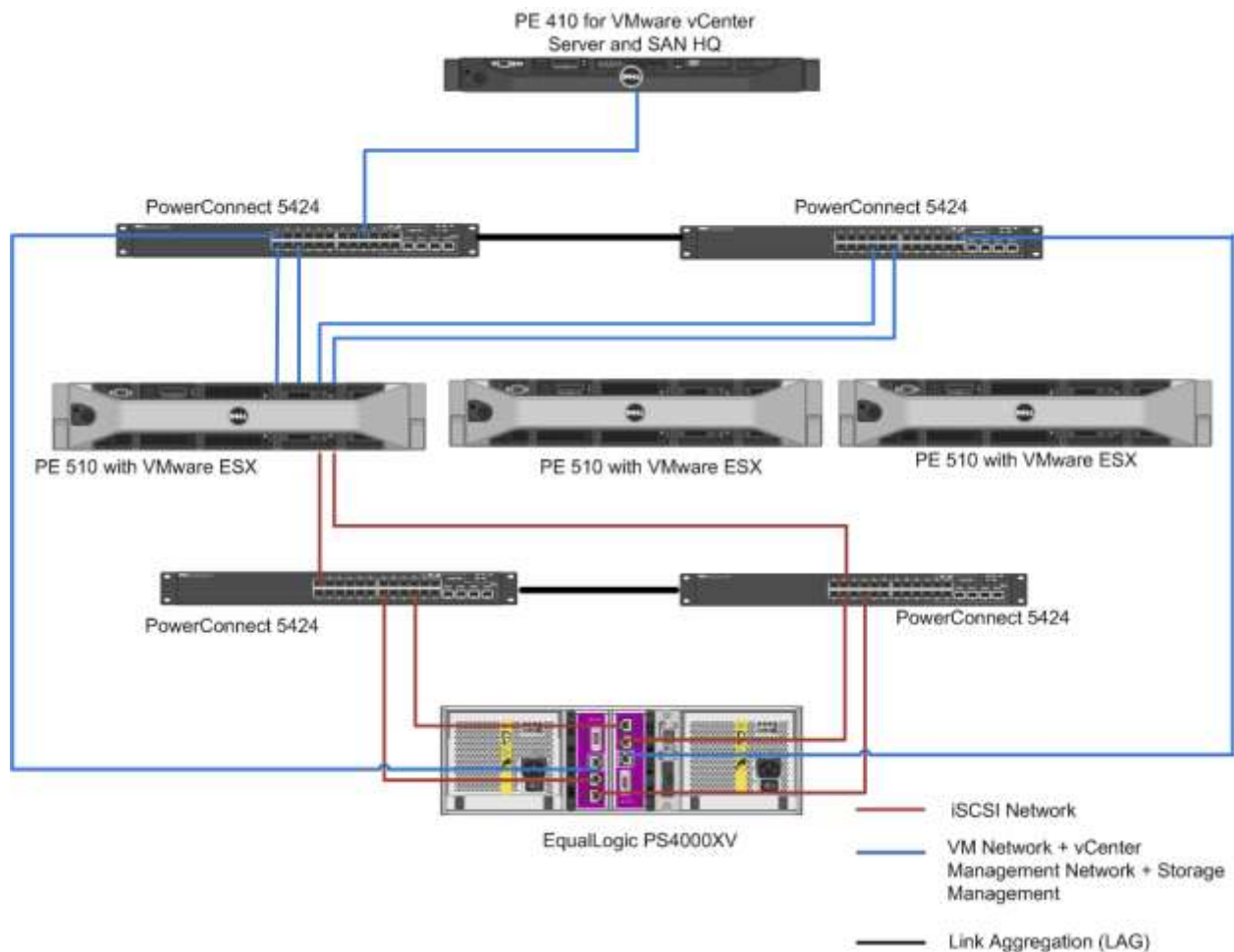
Figure 4 - Reference Architecture for High Availability + VMotion Configuration

Deploying an entry level virtual infrastructure is straight forward and hence the later sections of the document will focus on HA and HA + VMotion configuration.

6 Network Architecture

Typically a VMware vSphere infrastructure is comprised of four types of traffic: virtual machine traffic, management traffic, VMotion traffic and iSCSI traffic. In this section, we describe how the physical network adapters, PowerConnect switches, virtual network adapters and virtual switches are configured to support these various traffics and achieve load balancing and redundancy.

Both High Availability and High Availability + VMotion configurations, each host in the configuration has six network ports (Two LOMs+. Two dual port network adapters). For each PowerEdge R510 used in this configuration, the physical network adapters are enumerated as follows:

- *vmnic0, vmnic1*: First and second LOM ports on the 5716 dual port Gigabit Ethernet controller integrated on the system board.

- *vmnic2, vmnci3*: First and second ports, respectively, on the add-on first dual-port Broadcom NetXtreme II 5709 Gigabit Ethernet controller
- *vmnic4, vmic5*: First and second ports, respectively, on the add-on second dual-port Broadcom NetXtreme II 5709 Gigabit Ethernet controller

To support different traffic types in a virtualized environment, you can use two or four Dell PowerConnect 5424 switches, depending on the configuration being deployed. The PowerConnect switches are interlinked using a 4 port LAG (link aggregation), as shown in figures 3 and 4, which provides optimal bandwidth for the traffic between the switches and failover capabilities if there is a link failure. The PowerConnect supports a maximum of and up to eight Gigabit ports can be combined into a LAG, providing an aggregated bandwidth of 8 Gbps.

In the case of High Availability configuration, all network traffics types use the two physical switches and are isolated using VLANs. In the case of High Availability + VMotion configuration, four switches are used - two switches are dedicated for iSCSI SAN network and two switches are used only for the virtual machine network, management network and VMotion network, separated by VLAN. Figure 5 and 6 below shows how the network connections are done for a single ESX host in a High Availability and High Availability + VMotion configuration.
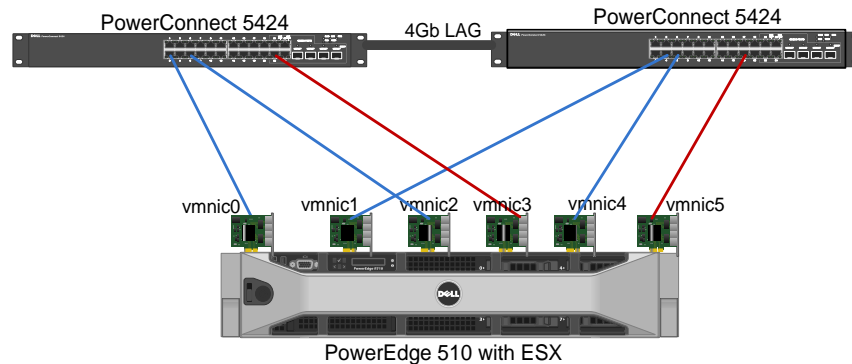

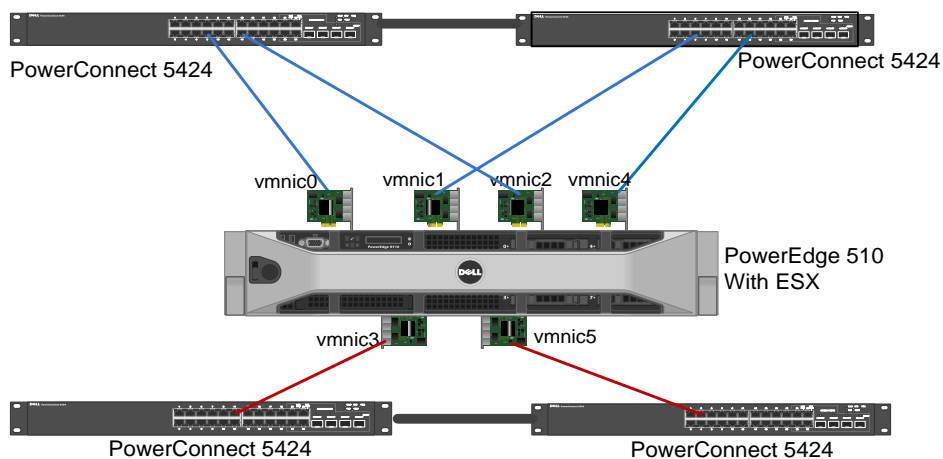**Figure 5: Network Configuration in ESX Host – High Availability Configuration**

Figure 6: Network Configuration in ESX Host – High Availability + VMotion configuration

6.1 Virtual Machine, Management and VMotion Network

For the virtual machine network, management network and VMotion network, four 1Gb Ethernet physical network adapters on the ESX host are teamed at the virtual switch. A virtual switch *vSwitch0* is created using these four network adapters as uplinks. Physically, these adapters on the host are connected to two physical network switches as shown in Figures 5 and 6. Three portgroups are created on *vSwicth0*: the first for VM network, the second for VMotion network and the third for host management or service console (for ESX host). Figure 7 illustrates the virtual switch configuration with port groups and how the virtual switch is connected to physical network adapters on the ESX host as shown in a vCenter interface.
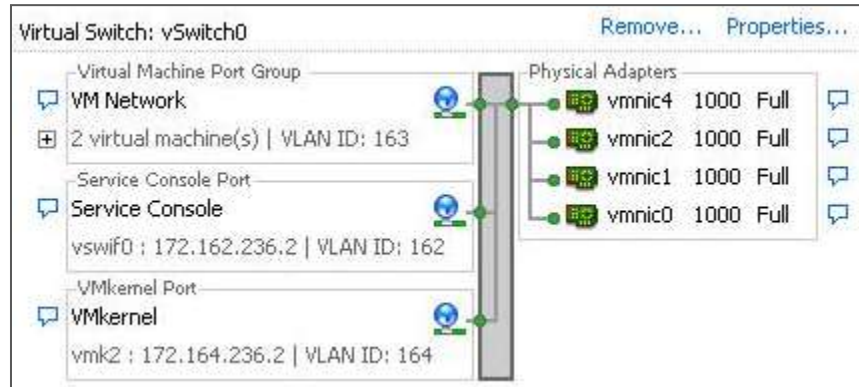

Figure 7: Portgroups as shown in vCenter

6.1.1 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.

Trunking must be used so that all the VLANs can share the same physical connection. To achieve this, all of the ports connected to ESX hosts on the physical switches are configured in the trunk mode.

VMotion traffic is unencrypted, so it is highly recommended that you isolate VMotion traffic. Using the VLAN configuration described previously, traffic isolation is achieved between the various traffic types, including the VMotion traffic.

Routing between VLANs is dependent on the specific customer requirements and is not included in the reference architecture.

6.1.2 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 for outbound traffic:

- **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 outbound transmission 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.

14

- **Route based on IP hash**: Here, the physical adapter is selected for outbound transmission based on the hash of the source and destination IP address. Link Aggregation Control Protocol (LACP) or EtherChannel (on Cisco switches) can be used with this setting. The physical switch load balancing algorithm determines which physical adapter is selected to receive the inbound 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 independently of each other. In case of multi-switch configurations (as illustrated in Figures 3 and 4), the switches need to be stacked. Since the stacking option is not available on PowerConnect 5400 series switches, this policy is not used.

For detailed description on the above mentioned load balancing policies refer to *VMware Virtual Network Concepts*.

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

Failover/Failback: Failback option is configured for Virtual Machine Network and VMotion Network. This setup enables both high availability and load balancing. Note that enabling failback may cause temporary loss of network access to virtual machines which fail back after an adapter recovers from failure. If this is not a tolerable situation, disable failback.

For Management network port group failback is disabled to avoid any loss of management traffic during the time that it takes for the port to be ready for serving traffic. This avoids the triggering of any false host isolation (if using VMware HA) due to a temporary loss of the management network

Note: The VMotion feature is not available in VMware vSphere Essentials Plus Edition. Hence, creating a VMotion port group is applicable only to High Availability + VMotion Configuration that uses VMware vSphere Advanced Edition.

6.2   iSCSI SAN Architecture

The iSCSI SAN is comprised of iSCSI traffic to the Dell EqualLogic PS Series arrays. This section discusses the best practices for configuring the iSCSI SAN using the software iSCSI initiator in ESX. The software initiator in the guest OSes is also supported, but is outside the scope of this document.

High Availability (HA) Configuration: The two 1Gb Ethernet network ports used for iSCSI communication on the ESX host are physically connected to the same PowerConnect switches used for the rest of the traffic, but are separated by VLANs. Hence, all four of the different traffic types share the same physical switches, but they are isolated using VLANs. For iSCSI network on the ESX host, two virtual switches (*vSwitch1* & *vSwitch2*) are created and one network port is added to each vSwitch. As shown in Figure 8 below, two VMkernel port groups (*vmkernel-iscis1* & *vmkernel-iscsi2*) and a VMkernel IP interface per portgroup (*vmk0* and *vmk1*) are created. VLAN tagging is enabled at portgroup level. These portgroups are used by the software iSCSI initiator to connect to the PS Series SAN. The ESX software iSCSI initiator is configured to make use of the two VMkernel interfaces for iSCSI traffic. Using two VMkernel interfaces, the software iSCSI initiator will see two paths to the target volume on PS Series SAN. Round Robin path selection policy can be used to balance the load between the two physical NIC ports.

High Availability (HA) + VMotion Configuration: The only difference between HA and HA + VMotion configuration is that in the HA + VMotion configuration, the network adapters used for iSCSI traffic, on the ESX host, are connected to two separate PowerConnect 5424 switches dedicated for iSCSI traffic. Since the traffic is isolated physically, there is no need to configure VLANs for iSCSI traffic in this configuration. The portgroups *vmkernel-iscsi1* and *vmkernel-iscsi2* need not be enabled for VLAN tagging in this configuration.

In both configurations, a two port LAG is configured between both switches to ensure high availability and bandwidth. The number of ports added to the LAG is equivalent to the number of active data ports in the PS array group.

Figure 8 illustrates the virtual switch configuration with port groups and how the virtual switch is connected to the physical network adapters on the host.
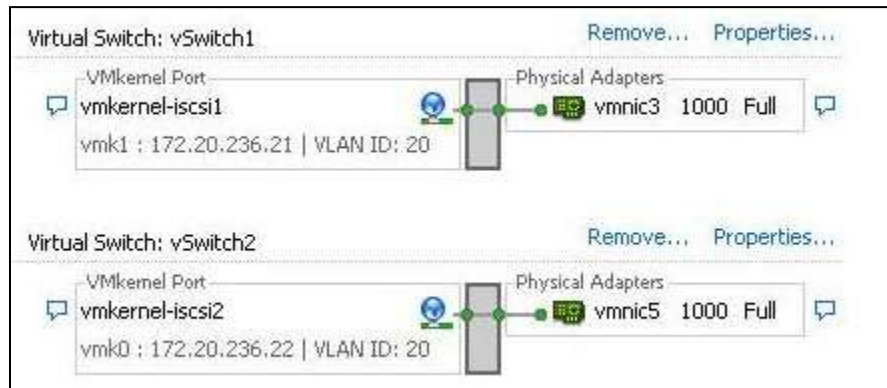

Figure 8: Configuring vSwitches and Portgroups for iSCSI

6.2.1   Storage Connectivity

Each controller in PS4000 has two 1Gb Ethernet ports for iSCSI and one 10/100 Mb port for management. As shown in Figure 9, the two 1Gb Ethernet ports for iSCSI are connected to the two switches (one from each controller) to each of the PowerConnect switches.

6.2.2   EqualLogic Network Requirements

In addition to the guidelines discussed previously, 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 (FastLink or Port Fast) 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.
- Enable Jumbo Frames on the switches, virtual switches and VMware kernel interfaces. See section 9 for details on this configuration.
- Disable iSCSI optimization on the PowerConnect 5424 switches used for iSCSI traffic.

7 Storage Configuration

The section provides recommendations for configuring the Storage Array.

Volume Size Considerations

Volume sizes depend on the customer environment and type of workloads. Using multiple NIC ports in conjunction with Round Robin path selection policy allows ESX to establish multiple connections to a single volume. Depending on the environment, you may decide to create multiple ~500 GB volumes with multiple VMs. It is important 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.

Load Balancing

Multipathing is a technique that allows more than one physical path to be used to transfer data between a host and an external storage device. Starting with ESX 4 or later, VMware provides the VMware Native Multipathing Plugin. This plugin supports three Path Selection Plugins; Most Recently Used (MRU), Fixed, and Round Robin (RR). The default selection for each volume in PS4000 is Fixed, but this can be changed to Round Robin to balance iSCSI traffic over multiple physical NIC ports and improve performance.

Dell EqualLogic SAN automatically distributes iSCSI connections between the available two network interfaces on each active controller. It also automatically distributes volumes across different storage arrays in the storage pool, if multiple storage arrays are present.

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 the 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 slightly lower performance and availability.

8 Management

This section covers the various components used to manage the solution.

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, 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.

System Management

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

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. **SAN HeadQuarters (SAN HQ)** is Dell's new management tool for consolidated performance and robust event monitoring across multiple virtualized EqualLogic SAN groups. It gathers and formats performance data and other vital group information for easy viewing. With SAN HQ, you have the ability to centrally monitor alerts, get historical performance reporting, do trend analysis for capacity planning and troubleshoot. For more information on Dell EqualLogic features and management capabilities, see www.dell.com/equallogic.

9   Deployment and Configuration Guidelines

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

9.1   Initial Setup

Rack, Stack, and Cable

1. Rack the Servers.
2. Rack the Dell EqualLogic storage.
3. Connect power to the servers, storage and switches.
4. Cable servers, storage and switches.
   a. Cable the server and storage for Ethernet connectivity as per the architecture.
      Interlink the switches.

Configuring the ESX Hosts

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

- Enable the CPU *Virtualization Technology* option available in the server BIOS.
- Configure RAID 1 on the internal hard drive.
- Deploy ESX on RAID 1 volume.
- Configure Service Console IP address on ESX host for management purpose.
- Configure Unified Server Configurator for self-update and firmware repository.

See the *ESX and vCenter Server Installation Guide* to install vCenter on the management server, http://www.vmware.com/pdf/vsphere4/r40/vsp_40_esx_vc_installation_guide.pdf

See the *Unified Server Guide*
http://www.dell.com/downloads/global/solutions/unified_server_overview.pdf

Configuring Network

Configure network switches. It is recommended that you create a configuration file for the network switches. These scripts are available for download here: http://www.delltechcenter.com/page/Virtualization+Scripts

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 PowerConnect Switch serial port. You will need a terminal program such as HyperTerminal.

TFTP: Configure an IP address for each PowerConnect Switch. **Telnet** or **ssh** can then be used to connect to the switch and **tftp** used to copy the `config` file to the switch. A TFTP server is required for this method.

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).

1. Create volumes for virtual machine storage. When creating a volume, using the EqualLogic group manager user interface that will be shared across ESX hosts, select the **Allow simultaneous connections from initiators with different IQN names** option.
   Create access lists for the volumes and add the IQN of all ESX hosts that will access the volumes. Ensure that you add the VMkernel iSCSI IPs to the access list. (You can use * to denote an IP address range.)

Configuring vCenter

Install VMware vCenter on the management server.
See the *ESX and vCenter Server Installation Guide* to install vCenter on the management server.
http://www.vmware.com/pdf/vsphere4/r40/vsp_40_esx_vc_installation_guide.pdf

9.2   Configuration Guidelines

1. Create a cluster in vCenter with VMware HA enabled. Add the ESX servers to the cluster.
2. For each server in the cluster, perform the following steps:

Create vSwitch0 (Management, Virtual Machine Network, and VMotion)

   a. Add vmnic0, vmnic1, vmnic2, vmnic4 to vSwitch0. Using vCenter GUI, under Configuration->Networking
   b. 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/Host Management VLAN 162, VM Traffic VLAN 163, and VMotion VLAN 164.
   c. Configure load balancing and failover (configure active/standby NIC) for each port group as per the guidelines provided in section 6.1.2.

Create vSwitch1 and vSwitch2 (For iSCSI)

Perform the following step-by step procedure to configure iSCSI software initiator in VMware vSphere to connect to PS series SAN in a multi-path environment.

   a. Using ESX host console, create new virtual switches vSwitch1 and vSwitch 2 for iSCSI enabling jumbo frame. Following commands below shows how to create vSwitch1 and enabling jumbo frame on it.

```
esxcfg-vswitch –a vSwitch1

esxcfg-vswitch –m 9000 vSwitch1

esxcfg-vswitch –a vSwitch2

esxcfg-vswitch –m 9000 vSwitch2
```

   b. Attach the NICs to vSwitch1 and vSwitch2 that will be used for the iSCSI network. To avoid single point of failure it recommended that you add vmnic ports in separate PCIe slots to vSwitch 1 and vSwitch 2. For example, add vmnic3 to vSwitch1 and vmnic5 to vSwitch2. Attaching the NIC can be done either through the CLI or using the GUI. Through the CLI, use the following commands:

```
esxcfg-vswitch –L vmnic3 vSwitch1

esxcfg-vswitch –L vmnic5 vSwitch2
```

c. Create a VMkernel portgroup and IP interface each for vSwitch1 and vSwitch2. Enable jumbo frames support for the VMkernel interface. Use the following commands to create VMkernel portgroups for iSCSI.

```
esxcfg-vswitch –A <portgroup name> <virtual switch>
```

```
esxcfg-vmknic –a –I <IP address> –n 255.255.255.0 –m 9000 <portgroup name>
```

For example, the following commands create a VMkernel portgroup named vmkernel-iscsi1 in vSwitch1 and assign an IP address enabling jumbo frame. In a similar way, vmkernel-iscsi2 VMkernel portgroup can also be created for vswitch2.

```
esxcfg-vswitch –A vmkernel-iscsi1 vswitch1
```

```
esxcfg-vmknic –a –I 172.20.236.21 –n 255.255.255.0 –m 9000 vmkernel-iscsi1
```

d. Enable VLAN tagging on the VMkernel port group used for iSCSI in the case of a two switch configuration.

```
esxcfg-vswitch –p vmkernel-iscsi1 –v 20 vSwitch1
```

```
esxcfg-vswitch –p vmkernel-iscsi2 –v 20 vSwitch2
```

For the procedure used to configure iSCSI SAN in a VMware environment, refer to the VMware's SCSI SAN configuration guide on the VMware Web site.

Configuring iSCSI Software Initiator for Multipathing

a. Enabling iSCSI initiator to connect to PS series SAN can be done through the GUI or using the CLI. To enable the iSCSI initiator through the CLI, run the following command on host service console.

```
esxcfg-swiscsi –e
```

b. Add the VMkernel ports created to the iSCSI initiator:
   i. Find the vmhba# from the vCenter GUI under *Configuration->Storage Adapters* or through the CLI using `esxcfg-scsidevs –s`
   ii. Determine the vmk# of each VMkernel port created.
   iii. Map each VMkernel port to the iSCSI software initiator.

```
esxcli swiscsi nic add –n <vmk#> –d <vmhba#>
```

For example, for the configuration shown in Figure 8, use the following commands

```
esxcli swiscsi nic add –n vmk0 –d vmhba35
```

```
esxcli swiscsi nic add –n vmk1 –d vmhba35
```

   iv. Verify that all of the vmk# are bound properly to the vmhba

```
esxcli swiscsi nic list –d <vmhba#>
```

Connect to PS series storage

a. To configure this, navigate in the vCenter GUI to Configuration -> Storage Adapters.
b. Click on the iSCSI Software Adapter and click **Properties**.
c. Click the Dynamic Discovery tab.
d. Click **Add**. In the iSCSI Server, give the Group IP Address of the PS Series SAN.
e. On performing a rescan of the storage adapters, under the iSCSI HBA adapters, the volumes created in the PS series storage will be detected by the host. There will be two paths to each volume detected by the host.

Enabling Multipathing – Round Robin Path Selection Policy (PSP)

To configure Round Robin PSP on a volume, select the volume from the vCenter GUI under Configuration -> Storage. Select the volume, right-click and choose the option Manage Paths. By default, fixed path will be chosen under Path Selection. Select Round Robin from the drop-down menu.

10 References

- Documentation for VMware vSphere 4 for Dell PowerEdge Servers
  http://support.dell.com/support/edocs/software/eslvmwre/VS/VS.htm
- A "Multivendor Post" on using iSCSI with VMware vSphere
  http://www.delltechcenter.com/page/A+%E2%80%9CMultivendor+Post%E2%80%9D+on+using+iSCSI+with+VMware+vSphere#
- VMware vSphere 4 iSCSI SAN Configuration Guide
  http://www.vmware.com/pdf/vsphere4/r40/vsp_40_iscsi_san_cfg.pdf
- Dell System Management Documentation (includes information for VMware ESX/ESXi)
  http://support.dell.com/support/systemsinfo/documentation.aspx?c=us&l=en&s=gen&~subcat=108&~cat=6
- 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
- VMware Network Concepts
  http://www.vmware.com/files/pdf/virtual_networking_concepts.pdf
- Configuring VMware vSphere Software iSCSI with Dell EqualLogic PS Series Storage
  http://www.equallogic.com/resourcecenter/assetview.aspx?id=8453