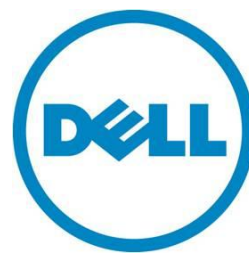

Implementing iSCSI SAN Best Practices with Dell Advanced Infrastructure Manager

This technical white paper explains how to separate SCN and iSCSI networks in AIM environment adhering to datacenter best practices.

Dell Global Solutions Engineering

May 2012



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

© 2012 Dell Inc. All rights reserved. Dell and its affiliates cannot be responsible for errors or omissions in typography or photography. Dell, the Dell logo, and PowerEdge are trademarks of Dell Inc. Intel and Xeon are registered trademarks of Intel Corporation in the U.S. and other countries. Microsoft, Windows, and Windows Server are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. 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.

May 2012 | Rev 1.0

Contents

Executive Summary	4
Introduction	5
About the iSCSI SAN Element	5
Creating an iSCSI SAN Element Inside AIM.....	5
Assigning the iSCSI SAN element to a persona	6
Server discovery with the iSCSI SAN element.....	7
Solution Architecture	8
Reference Architecture	10
Network architecture	12
Summary	14
References	15
Appendix	16
AIM Terminologies:	16

Tables

Table 1: Hardware configuration of PowerEdge R720 managed nodes	9
Table 2: Hardware configuration of PowerEdge R620 AIM controller and management server	9

Figures

Figure 1: Configuring the iSCSI SAN element in AIM web console	6
Figure 2: Assigning iSCSI SAN element to a persona.....	7
Figure 3: Persona with assigned iSCSI SAN element	7
Figure 4: Server NIC configuration inside the AIM console	8
Figure 5: Reference architecture diagram.....	11
Figure 6: iSCSI SAN basic network diagram	12

Executive Summary

Dell™ Advanced Infrastructure Manager (AIM) provides rapid workload mobility and recovery management across physical and virtual machines in a data center environment by seamlessly moving workloads and repurposing servers in minutes without network reconfiguration.

Data centers are typically designed for multiple networks that have different requirements for latency and resiliency. Deploying a separate, loss-intolerant network, dedicated to Internet Small Computer Systems Interface (iSCSI) storage is recommended as a best practice. AIM operates within a broad landscape and works closely with servers, storage, and network devices within the deployment framework. As a result, it is critical to architect an AIM deployment that adheres to storage and networking best practices to achieve efficiency, optimal performance, and security of the overall IT infrastructure.

In this context, AIM 3.5.1 architecture has introduced a feature called *iSCSI SAN element* that helps customers configure AIM architecture to comply with network configuration best practices recommended by storage vendors. With this feature, AIM supports hardware iSCSI booting for workloads using a dedicated iSCSI Storage Area Network (SAN) that is separate from the System Control Network (SCN). This network segregation helps to isolate iSCSI traffic from management traffic.

To provide more technical insight into this new feature, the Dell Global Solutions Engineering team built a reference architecture for AIM 3.5.1 using Dell PowerEdge™ 12th generation servers, Dell EqualLogic™ storage, and Dell PowerConnect™ switches. This white paper captures the architectural details, configuration guidelines, and best practices for using *iSCSI SAN Element* in an AIM 3.5.1 environment.

The key benefits from these networks segregation includes:

- Improved security by isolating SCN and iSCSI traffic.
- Efficient server repurposing by faster boot from persona.
- Better and efficient operations by using iSCSI specific configurations e.g. jumbo packets, priority traffic and flow control etc.

Introduction

In an enterprise data center, the heterogeneous server, storage, and networking infrastructures pose a challenge in terms of deployment, management, and overall life cycle management. Dell AIM solution helps customers manage heterogeneous environments efficiently. AIM leverages server images residing on shared SAN storage to enable faster server provisioning and workload mobility. AIM facilitates the creation of dynamic IT infrastructure where the server image is disassociated from the server hardware. The server image is stored in a centralized storage, and booted on a server that is dynamically selected for it. To accomplish this, network and storage access must be appropriately provisioned for the selected server.

Like any solution deployment, storage configuration for AIM architecture needs several considerations as recommended by storage vendors. Because AIM operations can be network intensive, an infrastructure that utilizes AIM should include network configurations for optimal storage performance. AIM 3.5.1 introduces a feature called the *iSCSI SAN element* that helps customers configure AIM architecture to comply with networking best practices as recommended for storage configurations.

For definitions of common AIM terminologies used in this paper, refer to the [AIM Terminologies](#): section in the Appendix.

For more details on Dell AIM see, <http://www.dell.com/aim>

About the iSCSI SAN Element

When deploying an iSCSI Storage Area Network (SAN), it is important to physically or logically separate storage traffic from ordinary LAN traffic to help the network run efficiently. Isolating storage traffic improves response times, prevents bottlenecks, minimizes potential performance issues, and enhances security. By ensuring the total segregation of SCN and iSCSI traffic, considerably better operational efficiencies are achieved by AIM 3.5.1, which is manifested in multiple ways such as enhanced data security & improved metrics of service level due to faster persona boot-up.

The iSCSI SAN element enables iSCSI storage to be located on an isolated iSCSI network. This differs from previous releases of AIM, in which the iSCSI storage responsible for storing personas can only reside on the System Control Network (SCN). The iSCSI SAN element acts as a DHCP server and is responsible for assigning IP addresses to the hardware iSCSI bootable Personas / VMRacks during boot time.

Prior to the iSCSI SAN element, the AIM controller required a network interface to be configured and wired to the iSCSI network. Multiple iSCSI SAN elements can also be configured to provide additional traffic isolation among different iSCSI networks, as long as the controller has multiple network interfaces configured and wired to iSCSI networks. In such cases, the controller should have multiple network interfaces configured and wired to the respective iSCSI networks

Creating an iSCSI SAN Element Inside AIM

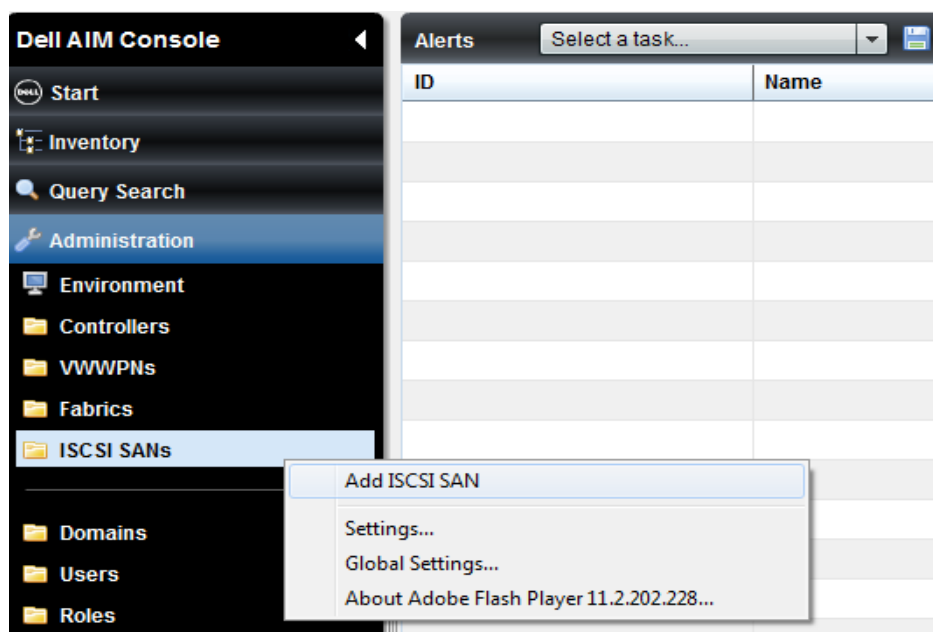
This section describes the sequence of events that occur while a persona boots on a dedicated iSCSI network with help from the iSCSI SAN element:

1. A controller request is made to start a persona on a server using hardware iSCSI booting.

2. The controller determines the storage target and target host's initiator. Then, it configures the iSCSI SAN element, which also acts as a DHCP server to provide an IP address in response to a request from server's iSCSI Network Interface Card (NIC).
3. The controller powers on the server from a management interface using Intelligent Platform Management Interface (IPMI) commands.
4. The server's iSCSI NIC makes a DHCP request, and fetches the target and initiator that will be used to boot the persona. The controller matches the NIC's Media Access Control (MAC) address to the binding set from step 2, and responds with the answer.
5. With the correct binding, the server attaches the storage LUN and boots the persona.

The iSCSI SAN element can be configured in the AIM web console or the AIM CLI. A service IP address is assigned to the iSCSI SAN element during its creation, and then the same service IP is assigned as a virtual IP on the controller's bonded interface. This bonded interface must be chosen as a service device during configuration.

Figure 1: Configuring the iSCSI SAN element in AIM web console



Additionally, a pool of IP addresses must be defined during the iSCSI SAN element creation. During server discovery and after persona boot up, the controller assigns IP addresses from this pool to the server's iSCSI NICs in response to DHCP requests.

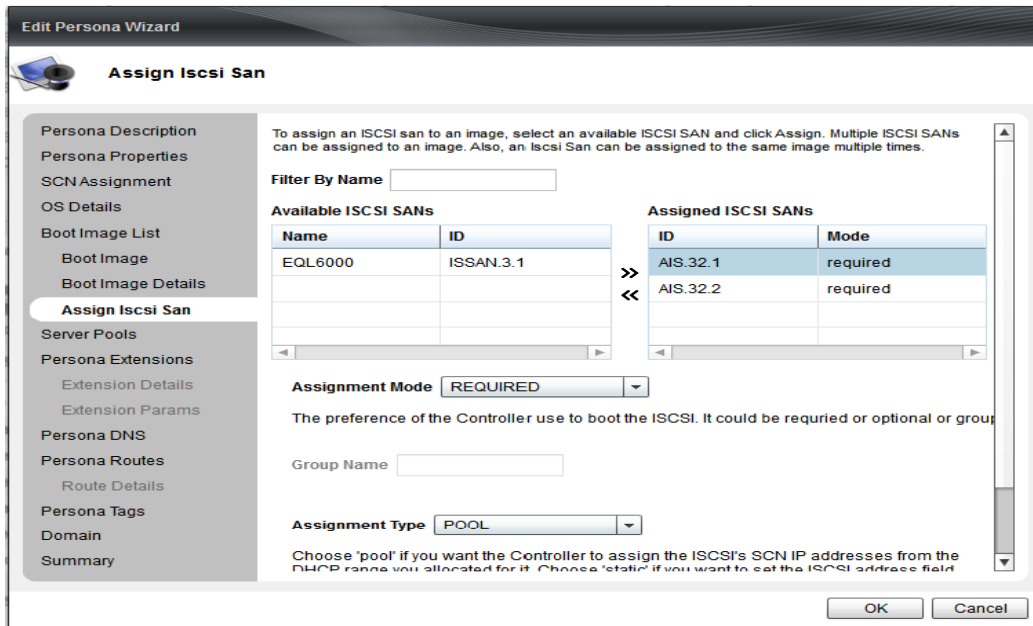
Assigning the iSCSI SAN element to a persona

After the iSCSI SAN element is configured, assign an iSCSI SAN element to personas and VM Racks to point to the images that are located on the storage array connected to the iSCSI network.

Assigning an iSCSI SAN element works in the same way as fabrics for Fibre Channel booting. An iSCSI SAN can be assigned to the same image multiple times, allowing a persona to boot only on a server that has at least as many iSCSI NICs as the number of iSCSI SAN elements assigned. Additional parameters can be set with this assignment from an assignment mode. There are three assignment modes available with AIM 3.5.1: Required, Optional, and Group.

Figure 22 shows one instance of the iSCSI SAN element assigned twice to a persona. This persona can now be booted only on a server with two or more iSCSI NICs.

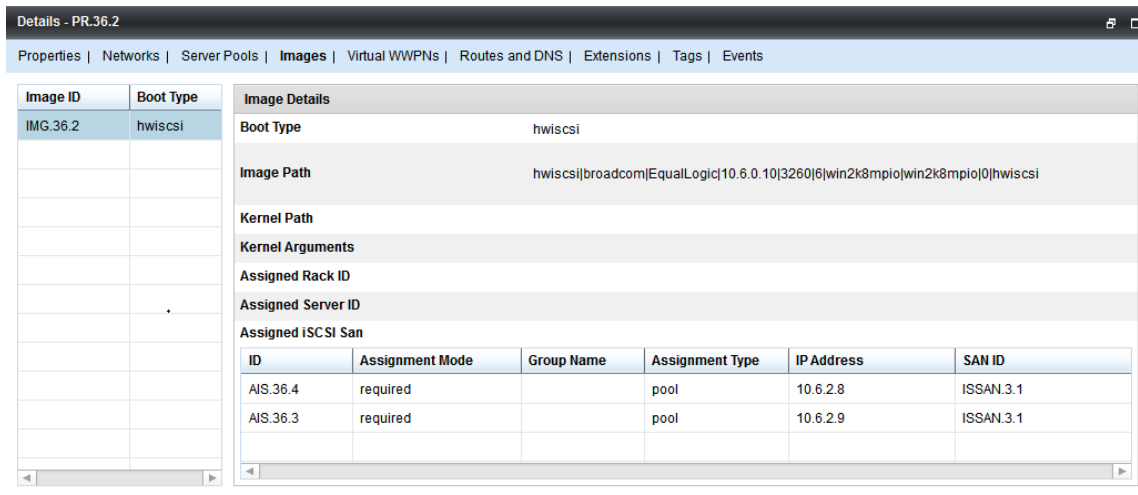
Figure 2: Assigning iSCSI SAN element to a persona



The *Required* parameter allows the AIM controller to boot a persona only on those servers where the required number of iSCSI NICs are present. The *Optional* parameter provides flexibility to boot a persona with fewer iSCSI NICs than the number of iSCSI elements assigned to the persona.

Figure 3 shows persona image properties after the iSCSI SAN element is configured and assigned.

Figure 3: Persona with assigned iSCSI SAN element



Server discovery with the iSCSI SAN element

The most common way to discover a new server is to configure the server to boot from its network interface card or PXE boot. A typical network-booted server can be used to host NFS, SAN, and software and hardware iSCSI-booted personas and VM Racks.

During discovery, the server is booted from the NIC or PXE booted with an image from the AIM controller, which discovers the server, including information about its serial number, CPU count, memory size, NICs with associated MAC addresses, HBAs with associated physical WWPNs, and so on. This information is sent back to the controller, which then shuts down the server using IPMI. The server remains powered off until a persona is assigned and started.

With AIM 3.5.1, a new server identification method has been introduced. The iSCSI SAN element is now responsible for server discovery in the AIM environment. This method is used only for personas that are hardware iSCSI booted. The NIC's hardware iSCSI initiator is used to boot hardware iSCSI booted personas. With this approach, discovery RAM disk can either reside on an iSCSI target server hosted by the AIM controller or on an EqualLogic storage array. The AIM controller mounts this discovery RAM disk as a LUN to the server's iSCSI NIC, and boots using LUN discovery. Once a server is discovered, its NICs are also discovered.

The NICs that are connected to an SCN on managed switches are assigned switch IDs and port IDs; however, NICs that are connected to the iSCSI network are not assigned IDs in AIM because iSCSI switches are not managed by the AIM controller. To configure iSCSI NICs on the iSCSI network, these NICs must be configured for hardware iSCSI booting on the server. Also, an iSCSI SAN element must be manually assigned to iSCSI NICs, along with the NIC role being set to *HWISCSIBOOT*.

After the NICs are configured, the discovery process is executed again to obtain the iSCSI MAC addresses. Alternatively, iSCSI MAC address can also be manually populated. Figure 4 shows the server NICs after discovery by the AIM controller. In Figure 4, the first four NICs are responsible for passing SCN and application traffic, and the remaining two NICs are responsible for passing iSCSI traffic.

Figure 4: Server NIC configuration inside the AIM console

Properties NICs HBAs Persona Assignment Server Pool Assignment Tags Events							
State	MAC Address	Channel	Role	Switch ID	Switch Port	iSCSI Mac Address ▲	iSCSI SA
Unknown	00-10-18-84-2F-DC	1	managed	SW.1.2	1/g37	00-10-18-84-2F-DD	
Unknown	00-10-18-84-2F-DE	1	managed	SW.1.2	1/g38	00-10-18-84-2F-DF	
Unknown	14-FE-B5-D9-54-D7	1	boot	SW.1.2	1/g33	14-FE-B5-D9-54-D8	
Unknown	14-FE-B5-D9-54-D9	2	boot	SW.1.2	1/g34	14-FE-B5-D9-54-DA	
Unknown	14-FE-B5-D9-54-DB	0	hwiscsiboot			14-FE-B5-D9-54-DC	ISSAN.3.1
Unknown	14-FE-B5-D9-54-DD	0	hwiscsiboot			14-FE-B5-D9-54-DE	ISSAN.3.1

Solution Architecture

- Three Dell PowerEdge R720 servers were used as managed nodes in the solution architecture.
- Two PowerEdge R620 servers were used: one for deploying the AIM controller and one as a management server for AIM and the virtual environment. The management server is an optional component that can be substituted by any server with access to the AIM environment.
- Two PowerConnect 6248 switches were used for the SCN network.
- Two PowerConnect 8024F switches were used for the iSCSI network.

- One EqualLogic PS6110XV iSCSI storage array was used for storing personas, VM Racks and user data.

The detailed hardware configuration for managed nodes, the AIM controller, and the management server is shown in Table 1 and Table 2.

Table 1: Hardware configuration of PowerEdge R720 managed nodes

	Components	Details
Processor	Intel® Xeon® E5-2660 2.2GHz 8-core, 20MB L3 Cache, Turbo, Hyper-Threading	Two Intel Xeon E5-2660s are populated in the two sockets
Memory	24GB (6 x 4GB) RDIMMS	DDR3, 1333 MHz or 1600 MHz Dual-Ranked RDIMMS
Network Controller	Broadcom 5720 rNDC Ethernet Controller	rNDC provides four 1Gb Ethernet ports
	1 x Broadcom 57810 Dual-Port 10Gb Network Controller	Add-in NIC which provides two 10Gb Ethernet ports
Persona Operating System	Microsoft® Windows Server® 2008 R2 SP1	Enterprise Edition
Out-of-band Management	iDRAC7 Enterprise	For remote management of the managed nodes

Table 2: Hardware configuration of PowerEdge R620 AIM controller and management server

	Components	Details
Processor	Intel Xeon E5-2660 2.2GHz 8-core, 20MB L3 Cache, Turbo, Hyper-Threading	Two Intel Xeon E5-2660s are populated in the two sockets
Memory	24 GB (6 X 4GB, DDR3)	DDR3, 1333 MHz
Network Controller	Broadcom 5720 rNDC Ethernet Controller	rNDC provides four 1Gb Ethernet ports
AIM Controller OS	RHEL 5.5	Enterprise Edition
Management Server OS	Windows Server 2008 R2	Standard Edition

NOTE: Applications residing on top of personas may require prerequisites such as Active Directory Domain Services (AD DS) and a Domain Name Server (DNS). These components are not a part of the AIM infrastructure and not included in the reference architecture. Customers implementing applications on top of personas must have these components preinstalled in the environment.

PowerEdge R720 Server

The PowerEdge R720 uses the new Intel Xeon E5-2600 processor product family. The processor’s Intel Integrated I/O mechanism provides up to 80 PCIe lanes per server, and supports the PCIe 3.0 specification. Furthermore, the R720 has highly expandable memory: 24 memory slots with up to 32GB per DIMM, for a total memory capacity of 768 GB.

For more details, see <http://www.dell.com/us/enterprise/p/poweredge-r720/pd>

PowerEdge R620 Server

The R620 uses the new Intel Xeon E5-2600 processor product family. As a hyper-dense, two-socket, 1U rack server, the Dell PowerEdge R620 has a large memory footprint and highly-expandable memory: 24 memory slots with up to 32GB per DIMM, for a total memory capacity of 768 GB.

For more details, see <http://www.dell.com/us/enterprise/p/poweredge-r620/pd>

EqualLogic PS6110XV Storage Array

The Dell EqualLogic PS6110XV is a virtualized 10GbE iSCSI SAN that combines intelligence and automation with fault tolerance to provide simplified administration, enterprise performance and reliability, and seamless scalability. With 15,000 RPM SAS disk drives, the PS6110XV array provides high transactional and sequential performance with up to 7.2 TB of storage. The PS6110 line features two 10GbE controllers with SFP+ and 10GBASE-T port options.

For more details refer <http://www.dell.com/us/business/p/equallogic-ps6110xv/pd>

PowerConnect 6248

Dell PowerConnect 6248 has advanced switching capabilities, including Power over Ethernet (PoE), high-density, high-performance stacking, and 10 Gigabit Ethernet (GbE) capabilities. The 48 built-in copper GbE ports maximize server and workstation connectivity.

For more details, see <http://www.dell.com/us/enterprise/p/powerconnect-6248/pd>

PowerConnect 8024F

PowerConnect 8024F switches are high-density, 10 GB Ethernet switches designed for data center, aggregation, and unified fabric deployments requiring high throughput and availability. These high density 24-port 10 GB switches support virtualization, iSCSI storage, and 10 GB traffic aggregation, and provide full routing features. To simplify connectivity with EqualLogic arrays, the iSCSI auto-configuration feature automatically detects arrays and configures the switch for optimal data transfer. This feature is enabled by default.

For more details, see <http://www.dell.com/us/enterprise/p/powerconnect-8024f/pd>

Reference Architecture

Figure 55 shows the reference architecture of Dell AIM 3.5.1 with segregated SCN and iSCSI networks.

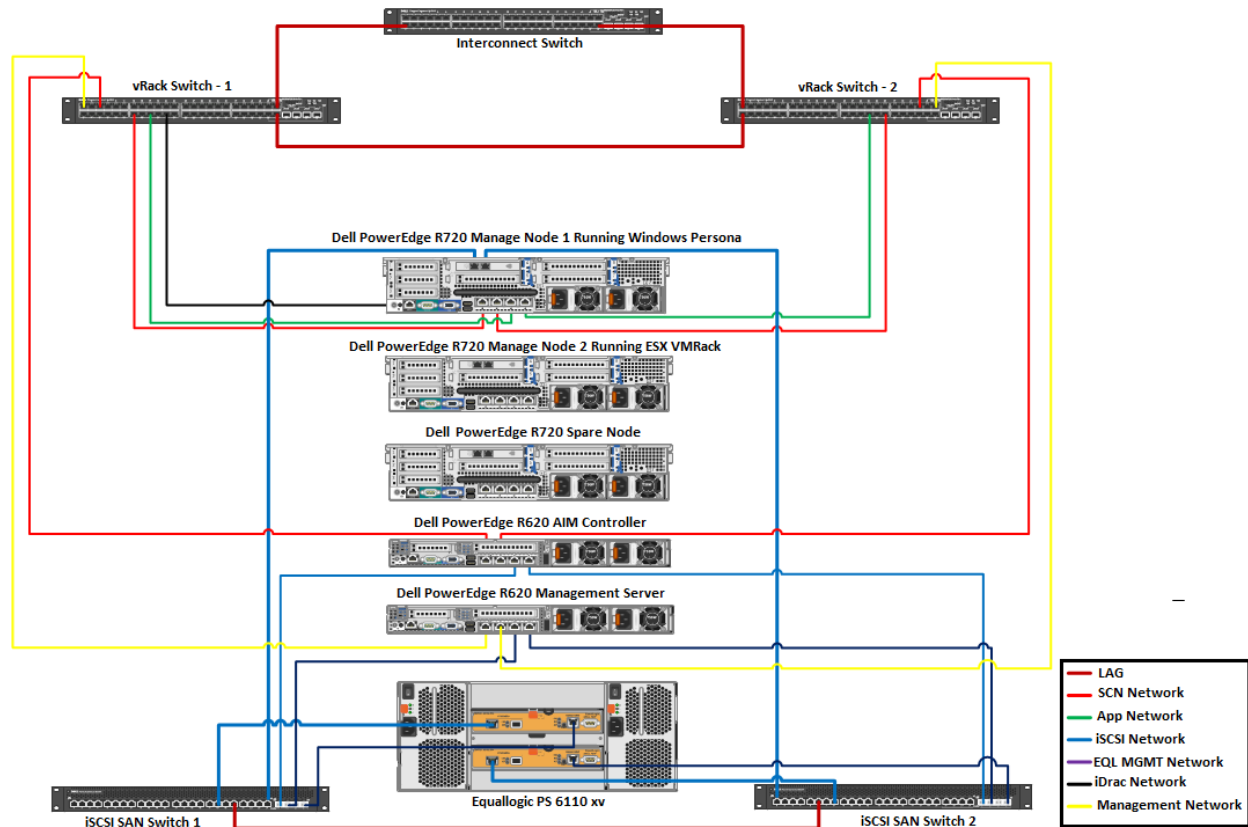
The Dell AIM controller was deployed on a PowerEdge R620 server with Active-Backup team mode to create two bonded interfaces to enable fault tolerance at the NIC level. The first bonded interface was used to pass SCN traffic, and the second bonded interface was used to pass iSCSI traffic. iSCSI offload mode was enabled on the iSCSI NICs of managed nodes.

A PowerEdge R620 server with Windows Server 2008 R2 Standard Edition was used as the management server responsible for managing the AIM and VMware environments and an EqualLogic PS6110 storage array. VMware vCenter Server 4.1 was also installed on this management server for managing VMware ESX VM Rack.

NOTE: Installation of AIM controller is beyond the scope of this document. For more details, see the product documentation.

The AIM reference architecture shown in Figure 5 has no single point of failure, and redundancy was incorporated into all components of the infrastructure.

Figure 5: Reference architecture diagram



As shown in Figure 55, A Windows Server 2008 R2 persona was created and started on the Dell PowerEdge R720 Manage Node 1. EqualLogic Multipathing¹ drivers were installed inside this Windows Server 2008 R2 persona for enabling fault tolerance at the NIC level.

Similarly, ESXi 4.1 VM Rack was created on Dell PowerEdge R720 Manage Node 2. Network connectivity to Manage Node 2 and a spare server are similar to Manage Node 1 (see Figure 54). In case of hardware failure, AIM retargets the failed persona and/or VM Rack onto an available spare server.

For more details on Dell EqualLogic iSCSI Optimization for Dell PowerConnect switches, see www.dell.com/downloads/global/products/pwcnt/Dell_EqualLogic_%20iSCSI_Optimization_for_Dell_Power_onnect_%20Switches.pdf

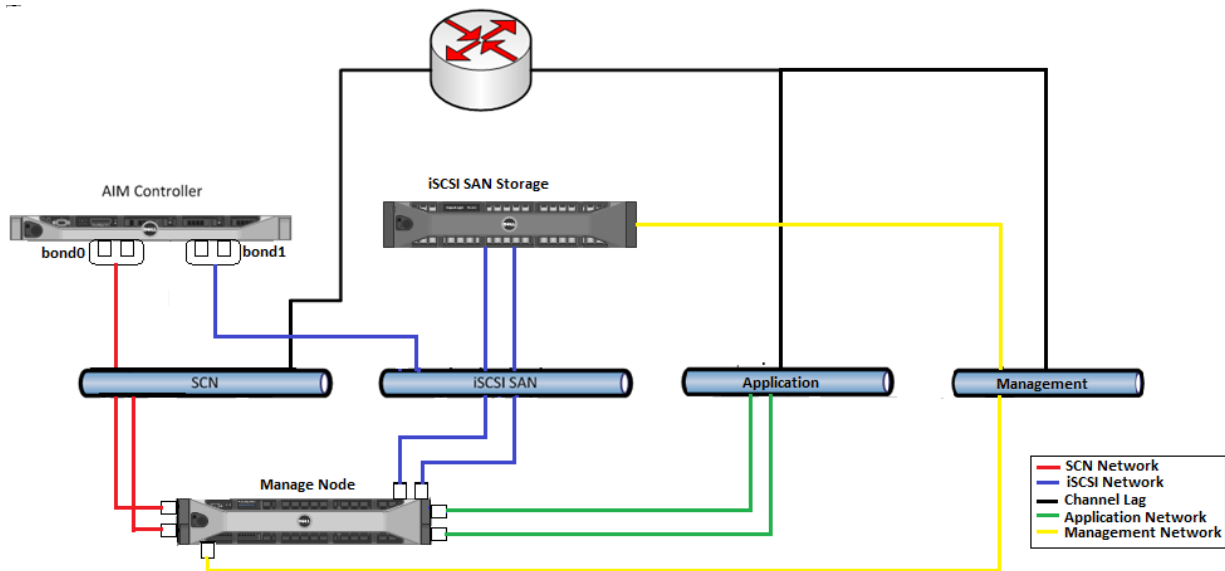
For more details on EqualLogic PS Series Best Practices for installing MPIO, see <http://support.equallogic.com/resourcecenter/assetview.aspx?id=5255>

NOTE: Creating a hardware-iSCSI bootable Windows persona and VM Rack are beyond scope of this document. For more details, see the product documentation.

Network architecture

Two bonded-interface with Active-Backup team modes were created on the AIM controller. The first interface (bond0) was comprised of LOM1 and LOM2, configured with an IP address on the SCN network, and wired to a switch port with Layer 3 connectivity to the SCN network. Similarly, the second bond interface (bond1) was comprised of LOM3 and LOM 4, configured with an IP on the iSCSI network, and wired to a switch port with Layer 2 connectivity to the iSCSI network. Bond1 was used as a service device during the iSCSI SAN element creation.

Figure 6: iSCSI SAN basic network diagram



The AIM infrastructure consists of two separate networks:

- **LAN** - Supports management network, application network, and the SCN. In addition, uplink to interconnect switch provides connectivity to solution support services (AD, DNS, DHCP, and so on).
- **SAN** - Supports the iSCSI network.

LAN architecture

As shown in Figure 55, the LAN includes two PowerConnect 6248 switches (vRack Switch 1 and vRack switch 2) to support the SCN, application network and management network. These traffic types were logically separated through VLANs.

A Channel Link Aggregation Group (LAG) was created between the two vRack switches and the interconnect switch to provide fault tolerance. In addition, uplinks to the interconnect switch connected AIM environment to solution support services (AD, DNS, NTP, and database for management applications).

The two vRack switches were fully managed by AIM; the AIM controller could make any necessary change to the switch configuration files in response to networking requirements. As a result, the host ports were configured by the controller with the SCN VLAN.

The traffic on the two vRack switches were segregated into three VLANs:

- SCN network (VLAN 4004)
- Application network (VLAN 6)
- Management network (VLAN 5)

Application networks or other networks can be connected to the personas using the AIM web console. When the persona is assigned and booted on a server, the network settings specified by Network Connections are configured transparently on either the physical interface in the server or on tagged interfaces (vNICs) created on top of the physical interface by the AIM agent.

SAN architecture

The SAN architecture in this solution includes two PowerConnect 8024F switches (iSCSI switch 1 and iSCSI switch 2) that support iSCSI network. These two iSCSI switches were connected using a channel LAG for providing fault tolerance. They were not discovered in the AIM environment, and therefore kept unmanaged. The EqualLogic PS 6110XV storage array was connected to both switches. The traffic on these two iSCSI switches was segregated into single VLANs:

- iSCSI network (VLAN 8)

The iSCSI SAN switches used in this architecture were configured while keeping certain best practices in mind, such as:

- A separate subnet and VLAN for the IP storage network was configured so that dedicated bandwidth and storage traffic can be contained within storage network.
- Access controls and traffic filtering were configured to ensure only authorized devices can access storage or administrative traffic.
- Individual ports connected to the storage nodes and host servers were kept in auto negotiate speed, duplex, and flow control at the host/node port level.
- Flow control was implemented on the storage network switch infrastructure. Failure to implement flow control at the server, on all storage and network ports connected to the SAN, can cause storage performance issues.
- The network environment within AIM has been configured to ensure that the storage environment can access network services through direct connections or appropriate routing.
- For building a high-availability IP storage network, multiple iSCSI switches were typically connected into a Single Layer 2 broadcast domain using multiple interconnects (Channel LAG). To avoid Layer 2 loops, the Spanning Tree protocol in portfast mode was implemented in the

switch infrastructure. Failing to do so can cause numerous issues on the IP storage networks, including performance degradation and traffic storms.

- Jumbo frame support was enabled on the two iSCSI switches, SAN storage array, and iSCSI NICs of all servers connected to the SAN.

NOTE: Tagged VLANs may also be used on the controller to access additional iSCSI networks if there are limited physical interfaces available. If VLAN tagging is enabled on the switch, then prior to booting the hardware-iSCSI bootable persona, the relevant VLAN tag must be assigned on the host NIC through the NIC configuration utility.

Summary

This paper provides reference architecture for AIM 3.5.1, and introduces the iSCSI SAN element feature. This new feature enhances AIM deployments by providing the ability to isolate SCNs from iSCSI networks, so that SCN network traffic does not interfere with iSCSI network traffic. It also enables AIM to support hardware iSCSI-booting for workloads using a dedicated iSCSI SAN network. This enables AIM personas to boot more quickly, improving operational efficiency. To summarize, AIM 3.5.1 helps in deploying efficient, secured and robust architecture for managing data center infrastructure.

References

Dell AIM:

<http://www.dell.com/us/enterprise/p/dell-advanced-infrastructure-manager/pd>

Other Dell AIM White papers:

<http://en.community.dell.com/techcenter/virtual-integrated-system/w/wiki/dell-aim-tech-whitepapers.aspx>

Dell EqualLogic iSCSI optimization for PowerConnect switches:

http://www.dell.com/downloads/global/products/pwcnt/Dell_EqualLogic_%20iSCSI_Optimization_for_Dell_Power_onnect_%20Switches.pdf

EqualLogic PS Series Best Practices for installing MPIO:

<http://support.equallogic.com/resourcecenter/assetview.aspx?id=5255>

iSCSI Protocol: Broadcom NetXtreme II Network Adapter User Guide:

<http://support.dell.com/support/edocs/network/BroadCom/R230837/en/iscsi.htm>

EqualLogic:

<http://www.dell.com/EqualLogic>

PowerConnect Switches:

<http://www.dell.com/us/enterprise/p/switch-powerconnect>

Appendix

AIM Terminologies:

This section introduces common terminologies used in the AIM environment.

- **AIM Controller:** AIM controller is the heart of the AIM solution which manages the entire infrastructure environment. The AIM Controller is software that manages the physical and virtual hardware, software, and network configurations in the AIM environment. It runs on a dedicated server or on ESX VM outside the AIM environment, communicating with the environment via its existing network infrastructure.
- **AIM Console:** The web-based user interface, AIM Console is used to monitor and work with the AIM environment. It is hosted on the AIM Controller.
- **Persona** - A server environment captured on disk that includes the OS, optional AIM agent software, application software, and network and other settings. It is typically stored on a central storage repository. Personas can reside on a server's local disk or on any of a number of kinds of Ethernet or Fiber Channel network storage resources, including NFS, NAS, iSCSI, and SAN-based storage servers.
- **Retarget** -It constitutes multiple steps - (i) Stopping the persona (ii) Assigning the persona to another server (iii) Booting the reassigned server onto the OS image corresponding to the persona (iv) Reconfiguring the persona and its connectivity.
- **VMRack** - A VMware ESX host or Microsoft Hyper-V host that is under AIM management and their associated virtual machines (VMs). AIM treats VMs in a VMRack as any other physical server; therefore configuring personas to boot on both physical servers and virtual machines becomes easier when needed.
- **SCN** - The System Control Network (SCN) is utilized by AIM to discover new servers and their capabilities by performing PXE for the initial part of the boot process, Controller and aim agent communication and as a first hop network used by the server NIC to connect to IP based central storage (NFS, iSCSI) to boot the Personas / VMRacks.