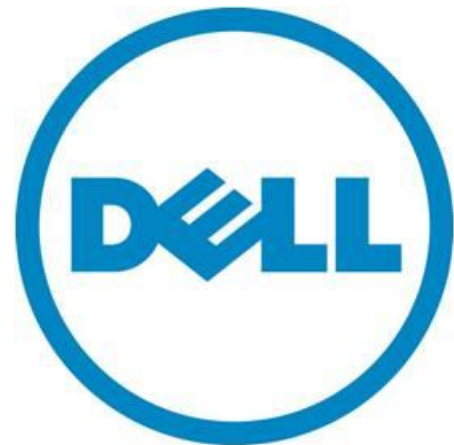


Business-Ready Configuration for Microsoft Hyper-V R2 on Dell PowerEdge M-Series Servers with EqualLogic Storage

A solution guide



Brian Gautreau

Dell Virtualization Solutions Engineering

<http://www.dell.com/virtualization/businessready>

Feedback: solutionfeedback@dell.com

THIS WHITE PAPER 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.

© 2010 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

Dell, the DELL logo, and the *DELL* badge, *PowerEdge*, *PowerConnect*, *EqualLogic*, and *OpenManage* are trademarks of Dell Inc. *Microsoft*, *Windows*, *Windows Server*, *Hyper-V*, and *Active Directory* are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. *Broadcom*, *BASP*, and *BACS* are trademarks or registered trademarks of the Broadcom Corporation. 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 Inc. disclaims any proprietary interest in trademarks and trade names other than its own.

November 2010

Table of Contents

Introduction	5
Audience and Scope	5
Overview	5
Dell PowerEdge M910	7
Dell PowerEdge M610	7
Dell EqualLogic PS6010X	7
Dell PowerConnect M8024	8
Specifications.....	8
Reference Architecture.....	10
Design Principles	10
Figure 1 - Architecture Overview	11
Network Architecture	11
Figure 2 - Network Fabric Overview	12
Data Networks, Subnets, and VLANs.....	12
Switch Trunking	13
Spanning Tree	13
Virtual Machine Network Fault Tolerance	13
Uplinks (Off-Rack Connectivity)	14
M910 Network Configuration.....	14
Figure 3 - M910 Network Connections.....	15
M610 Network Configuration.....	16
Figure 4 - M610 Network Connections.....	17
Storage Architecture	18
Figure 5 - Storage Volume Architecture	18
Shared-Storage Volume Size Considerations	18
Array RAID Configuration	19
Jumbo Frames	19
Scalability	20
Scaling Compute Resources	20
Scaling Storage Resources	20
Figure 6 - Scalable Option.....	21
Management	21
Physical Server Management.....	21
(a) Chassis Management Controller (CMC)	21

(b) Keyboard, Video, Mouse (KVM)	21
(c) iDRAC	22
Storage Management	22
(a) EqualLogic Group Manager	22
(b) SAN HeadQuarters.....	22
Systems Management	22
(a) Dell OpenManage	22
(b) Unified Server Configurator (USC)	22
Virtual Machine Management	23
Deployment	23
Network Configuration	23
Server Configuration	23
Storage Configuration	24
Failover Clustering Configuration	24
Management Configuration.....	25
How to Order the Reference Configuration	25
Dell Global Services	26
Additional Reading and Resources	27
Dell PowerEdge Server Documentation and Hardware/Software Updates	27
Dell PowerConnect Switch Documentation and Firmware Updates.....	27
Dell Virtualization Documentation.....	27
Microsoft® Hyper-V Documentation	27
Microsoft® Management Software	27
Broadcom Networking.....	28
Microsoft® Roles and Features	28
Appendix A: Reference Network Configuration	28
PowerConnect M8024 Switch 1 Configuration File Used for Host Management, Cluster, Live Migration, and Virtual Machine Traffic.....	28
PowerConnect M8024 Switch 2 Configuration File Used for Host Management, Cluster, Live Migration, and Virtual Machine Traffic.....	30
PowerConnect M8024 Switch Configurations Used for iSCSI Network	32

Introduction

Business-Ready Configurations for Virtualization are a reference architecture family of products designed to aid in the ordering, deployment, and maintenance of a virtualization infrastructure. These architectures are designed to meet specific customer needs through the use of various server, storage, interconnect, and virtualization technologies available from Dell.

The reference architectures defined in this document are targeted at enterprise business virtualization needs, although others may also benefit from this architecture. The solutions include Dell PowerEdge™ M610 and M910 servers, Dell EqualLogic™ PS6010X series storage, Dell PowerConnect™ switches, Microsoft Hyper-V® technology, and Microsoft System Center® management software. The extensive design and engineering work put into these solutions allow customers to quickly and confidently deploy these architectures into production environments, thereby helping to eliminate costly and time consuming trial and error work often encountered during complex deployments.

This guide provides information for both before and after the purchase of a Business-Ready Configuration. Prior to purchase, this information can aid in the solution sizing, licensing selection, and preparation of the solution deployment environment. After purchase, this guide can aid with the setup, configuration, and deployment of the solution.

Audience and Scope

The intended audiences for this white paper are IT administrators, IT managers, and channel partners who are planning to deploy or resell Microsoft Hyper-V virtualization solution for themselves or for their customers. The reference architectures provide an overview of the recommended servers, storage, software, and services. They can be used to plan, scope, and procure the required components to set up a virtualization infrastructure. It is assumed that the reader has a basic understanding of Hyper-V server virtualization, iSCSI storage, and networking concepts.

Overview

As mentioned, the reference architectures discussed in this white paper are centered on Microsoft's latest virtualization platform, Microsoft Windows Server® 2008 R2 with Hyper-V (referred to here on as Hyper-V R2), PowerEdge M610 and M910 servers, and Dell EqualLogic PS6010X storage, PowerConnect M8024 for iSCSI, public, and cluster networking, and Microsoft System Center Virtual Machine Manager 2008 R2 (SCVMM R2) which provides a solution for managing Hyper-V R2 virtual infrastructure. This section provides an overview of the hardware components used for the Business-Ready Configuration followed by the key capabilities it is offering.

The three reference architectures discussed in this white paper are small, medium, and large sized configurations. A high-level overview of the hardware included in each configuration is shown in Table 1.

Table 1 - Hardware Overview

	Small	Medium	Large
Virtualization Hosts	(2) x PowerEdge M910	(4) x PowerEdge M910	(6) x PowerEdge M910
External Storage	(1) x PS6010X (16 x 600GB 10K SAS disks)	(2) x PS6010X (32 x 600GB 10K SAS disks)	(3) x PS6010X (48 x 600GB 10K SAS disks)
Network Infrastructure	(4) x PowerConnect M8024	(4) x PowerConnect M8024	(4) x PowerConnect M8024
Virtualization Management	(2) x PowerEdge M610	(2) x PowerEdge M610	(2) x PowerEdge M610
Key features	<p>Ideal for deployment of up to 60 VMs.</p> <p>High Availability provided by Failover Clustering.</p> <p>iSCSI based storage array (up to 9.6TB of storage capacity).</p> <p>Highly available single console management using SCVMM 2008 R2.</p>	<p>Ideal for deployment of up to 120 VMs.</p> <p>High Availability provided by Failover Clustering.</p> <p>iSCSI based storage array (up to 19.2 TB of storage capacity).</p> <p>Highly available single console management using SCVMM 2008 R2.</p>	<p>Ideal for deployment of up to 180 VMs.</p> <p>High Availability provided by Failover Clustering.</p> <p>iSCSI based storage array (up to 28.8TB of storage capacity).</p> <p>Highly available single console management using SCVMM 2008 R2.</p>

NOTE: The VM counts listed here are based upon VMs with an average of 4 GB RAM, 50GB shared storage space, and one virtual CPU. Actual results will vary.

Dell PowerEdge Blade Servers



- Energy efficient PowerEdge M1000e enclosure
- 11th generation M910 blade server
 - Intel Xeon 7500 processor
 - Up to 512GB DDR3 Memory
 - FlexMem Bridge Technology
- 11th generation M610 blade server
 - Intel Xeon 5500 processor
 - Up to 192GB DDR3 Memory
- Flex Address
- CMC and iKVM for enclosure management



Dell EqualLogic PS6010 Series iSCSI arrays

- Intelligent and automated management tools
- Self-Managing Array
- Seamless scalability
- EqualLogic multi-pathing Device Specific Module for Microsoft iSCSI initiator integration

Dell PowerEdge M910

The PowerEdge M910 is a part of the 11th Generation PowerEdge portfolio. The M910 is a four socket, full height blade server supporting the Intel Xeon 7500 series processors, DDR3 memory, and the optional FlexMem Bridge enabling systems using two to access all 32 DIMMs. Without FlexMem, dual processor systems would be limited to 16 DIMMs. It is available with the modular PERC H200 or H700 RAID controllers and 2.5” 6Gbs SAS drives. The M910 has two embedded dual-port Broadcom BCM-5907S 1 Gbps Ethernet adapters. The M910 has four available mezzanine risers that enable the server to connect with the blade chassis’ I/O fabric. The options available for the M910 make it an ideal server for virtualization.

The PowerEdge M910 includes the Remote Management capabilities of the iDRAC6 Enterprise. The remote management capabilities provide an interface for administration tasks such as virtual media access, out-of-band alerting, status, and inventory.

Dell PowerEdge M610

The PowerEdge M610 server is also part of the 11th generation PowerEdge server portfolio. It is a half-height, two socket blade server supporting the Intel Xeon 5500 and 5600 series processors. It has a capacity of 12 DIMMs with a maximum of 192GB of DDR3 memory. The M610 uses the PERC on-board RAID controllers, along with the 2.5” 6Gbs SAS drives. For I/O Mezzanine connectivity, the M610 features two onboard BCM-5709S 1 Gbps Ethernet interfaces, and has two additional Mezzanine riser ports for the B and C IO fabrics. The VM and server density available with the M610 also make it a great server for virtualization environments.

Like the M910, the M610 server also features the iDRAC6 Enterprise out-of-band management interface. Enabling the system administrator to remotely power on/off, capture failed system messages, and connect virtual media.

Dell EqualLogic PS6010X

The Dell EqualLogic PS6010X is the virtualized iSCSI SAN storage device from Dell designed for deployment in a datacenter environment. It consists of dual 10 Gbps Ethernet network ports and one 10/100 Mb port dedicated for the management network plus 4GB of battery-backed memory per controller and options of 8 or 16, 600GB hot-pluggable SAS drives. The Dell EqualLogic storage arrays also provide the following capabilities:

Reliability: Dell EqualLogic PS6010X Series arrays have hot-swappable redundant components, a choice of RAID types 5, 6, 10, and 50, 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 operations per second, are increased. This increased capacity can be utilized without downtime. Thin-provisioning permits predefinition of a group of 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.

Dell PowerConnect M8024

The Dell PowerConnect M8024 delivers up to 24 ports (16 direct to blade plus 2 modular ports in any combination of 4-port SFP+, 3-port CX-4, or 2-port 10Gb Base-T RJ45) of 10 Gbps Ethernet. The M8024 shares many features, functions, and benefits of the entire PowerConnect switch series. It provides a full 240Gb/s switch fabric capacity, supports VLANs, QoS, and up to eight 10 Gbps Ethernet ports can be combined into a Link Aggregate Group (LAG), providing an aggregated bandwidth of 80 Gbps.

Specifications

Table 2 - Solution Components

	Small Solution	Medium Solution	Large Solution
Hardware			
Solution ID	1230400.1	1230423.1	1230428.1
Hyper-V R2 Server	(2) x M910	(4) x M910	(6) x M910
Storage Device	(1) x PS6010X	(2) x PS6010X	(3) x PS6010X
Management Servers	(2) x M610	(2) x M610	(2) x M610
Network Configuration			
Network Configuration	(4) x PowerConnect M8024 Network Switches	(4) x PowerConnect M8024 Network Switches	(4) x PowerConnect M8024 Network Switches

Hyper-V R2 Server Configuration			
Server Model	M910	M910	M910
Processor	(2) or (4) x Intel Xeon (Nehalem) L7555, 1.86Ghz, 24M Cache	(2) or (4) x Intel Xeon (Nehalem) L7555, 1.86Ghz, 24M Cache	(2) or (4) x Intel Xeon (Nehalem) L7555, 1.86Ghz, 24M Cache
Memory	128 GB (16 x 8GB, DDR3)	128 GB (16 x 8GB, DDR3)	128 GB (16 x 8GB, DDR3)
Onboard Network Adapter	(2) x Integrated Broadcom 5709S dual-port 1 Gbps Ethernet adapter	(2) x Integrated Broadcom 5709S dual-port 1 Gbps Ethernet adapter	(2) x Integrated Broadcom 5709S dual-port 1 Gbps Ethernet adapter
Add-in Controllers	(4) x Broadcom BCM57711 10 Gbps Ethernet Mezzanine Cards	(4) x Broadcom BCM57711 10 Gbps Ethernet Mezzanine Cards	(4) x Broadcom BCM57711 10 Gbps Ethernet Mezzanine Cards
Internal Storage Controller	PERC H200M	PERC H200M	PERC H200M
Internal Drive	(2) x 72 GB in RAID 1	(2) x 72 GB in RAID 1	(2) x 72 GB in RAID 1
Hypervisor	Microsoft Hyper-v R2 Enterprise Edition	Microsoft Hyper-v R2 Enterprise Edition	Microsoft Hyper-v R2 Enterprise Edition

External Storage Configuration			
Storage Device	(1) x PS6010X	(2) x PS6010X	(3) x PS6010X
Drives	(16) x 600GB, 10K SAS	(32) x 600GB, 10K SAS	(48) x 600GB, 10K SAS
Storage Capacity	9.6 Terabyte total drive capacity	19.2 Terabyte total drive capacity	28.8 Terabyte total drive capacity

Hyper-V R2 Management Server Configuration			
Server Model	M610	M610	M610
Processor	(2) x Intel Xeon (Nehalem) E5550, 2.66Ghz, 8M Cache	(2) x Intel Xeon (Nehalem) E5550, 2.66Ghz, 8M Cache	(2) x Intel Xeon (Nehalem) E5550, 2.66Ghz, 8M Cache
Memory	16 GB (4 x 4GB)	16 GB (4 x 4GB)	16 GB (4 x 4GB)
Network Controller	(2) x Integrated Broadcom 5709S dual-port 1 Gbps Ethernet adapter	(2) x Integrated Broadcom 5709S dual-port 1 Gbps Ethernet adapter	(2) x Integrated Broadcom 5709S dual-port 1 Gbps Ethernet adapter
Add-in Controllers	(2) x Broadcom BCM-57711 10 Gbps Ethernet Mezzanine Cards	(2) x Broadcom BCM-57711 10 Gbps Ethernet Mezzanine Cards	(2) x Broadcom BCM-57711 10 Gbps Ethernet Mezzanine Cards
Operating System	Microsoft Windows Server 2008 R2 Enterprise Edition	Microsoft Windows Server 2008 R2 Enterprise Edition	Microsoft Windows Server 2008 R2 Enterprise Edition

Virtualization Management Software	Microsoft System Center Virtual Machine Manager 2008 R2	Microsoft System Center Virtual Machine Manager 2008 R2	Microsoft System Center Virtual Machine Manager 2008 R2
Server Management	Dell OpenManage™ 6.3.0	Dell OpenManage 6.3.0	Dell OpenManage 6.3.0

Reference Architecture

The following sections discuss the reference configuration architecture and design. These sections lay the ground work for the deployment of the reference configuration.

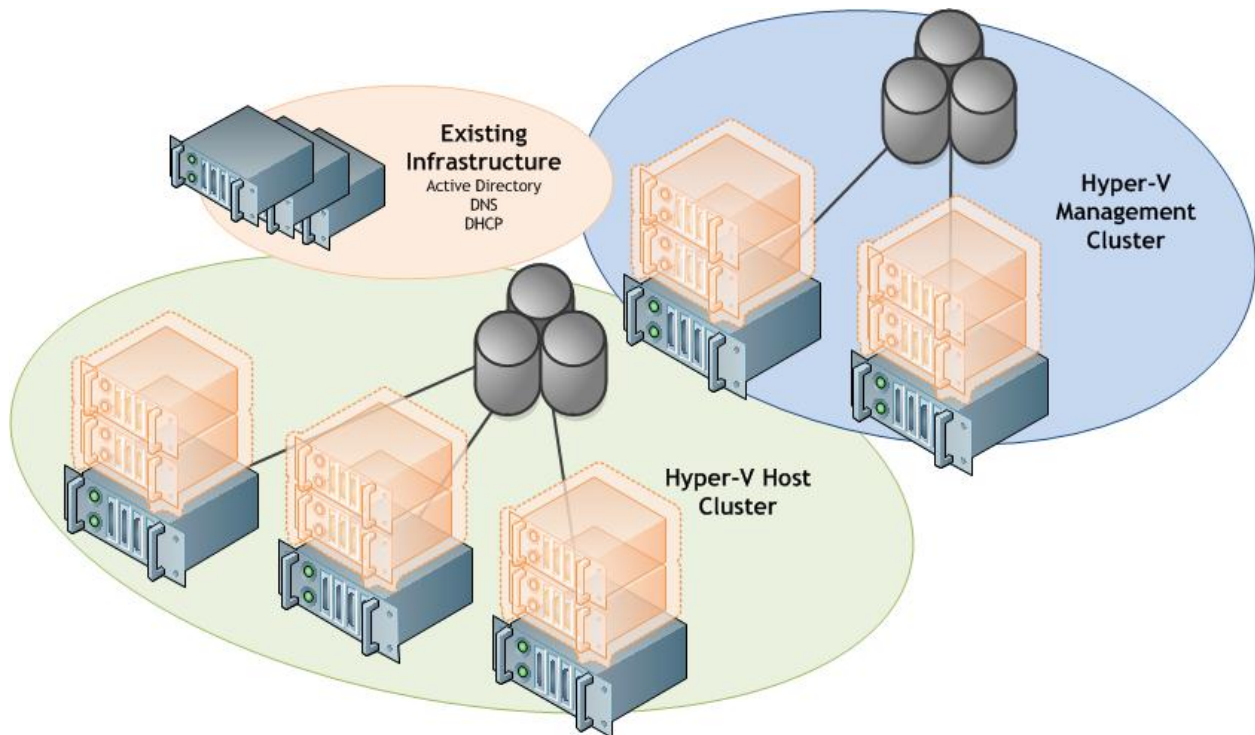
Design Principles

The following design principles were used while designing these configurations.

1. **Choosing optimal configuration:** The solution components for this Business-Ready Configuration are specifically designed for the requirements of a large virtualization environment. That being the case, the hardware components chosen are well suited for environments where VM density is a decision factor or VMs need a greater amount of memory and virtual CPUs. SCVMM R2 is chosen as the management software for its native Microsoft Hyper-V integration; extensibility with PowerShell and other Microsoft System Center products; and the number of servers being managed.
2. **High Availability and Live Migration Capabilities:** Configuring Hyper-V R2 hosts in Microsoft Failover Clustering enable VMs to move between hosts using live migration with minimal downtime. This also protects VMs from having a single point of failure. Management of the cluster is also made highly available through the use of Hyper-V R2 and Failover Clustering.
3. **Isolated and redundant network architectures:** The network configuration used in the reference architecture provides isolation for different networks in the virtual infrastructure and avoids a single point of failure.
4. **Virtualization Management Capabilities suited for Large Environments:** Management of the reference configuration discussed in this solution guide is by using Microsoft SCVMM R2, which is best suited for large deployments of Hyper-V and VMs. It supports management of up to 400 host operating systems environments and 8000 VM environments. It is a unified solution with a single console for managing both physical and virtual environments.
5. **Existing Infrastructure:** The design also assumes that environments will already be configured with core infrastructure components such as Microsoft Active Directory® (AD), Dynamic Host Configuration Protocol (DHCP), and Domain Name Service (DNS).

To provide highly available management of the configurations, two Hyper-V failover clusters are created. The first cluster is the Hyper-V management cluster and the second is the Hyper-V host cluster. This ensures that all features of the system as a whole are highly available (hosts and management). A reference configuration overview appears in Figure 1.

Figure 1 - Architecture Overview



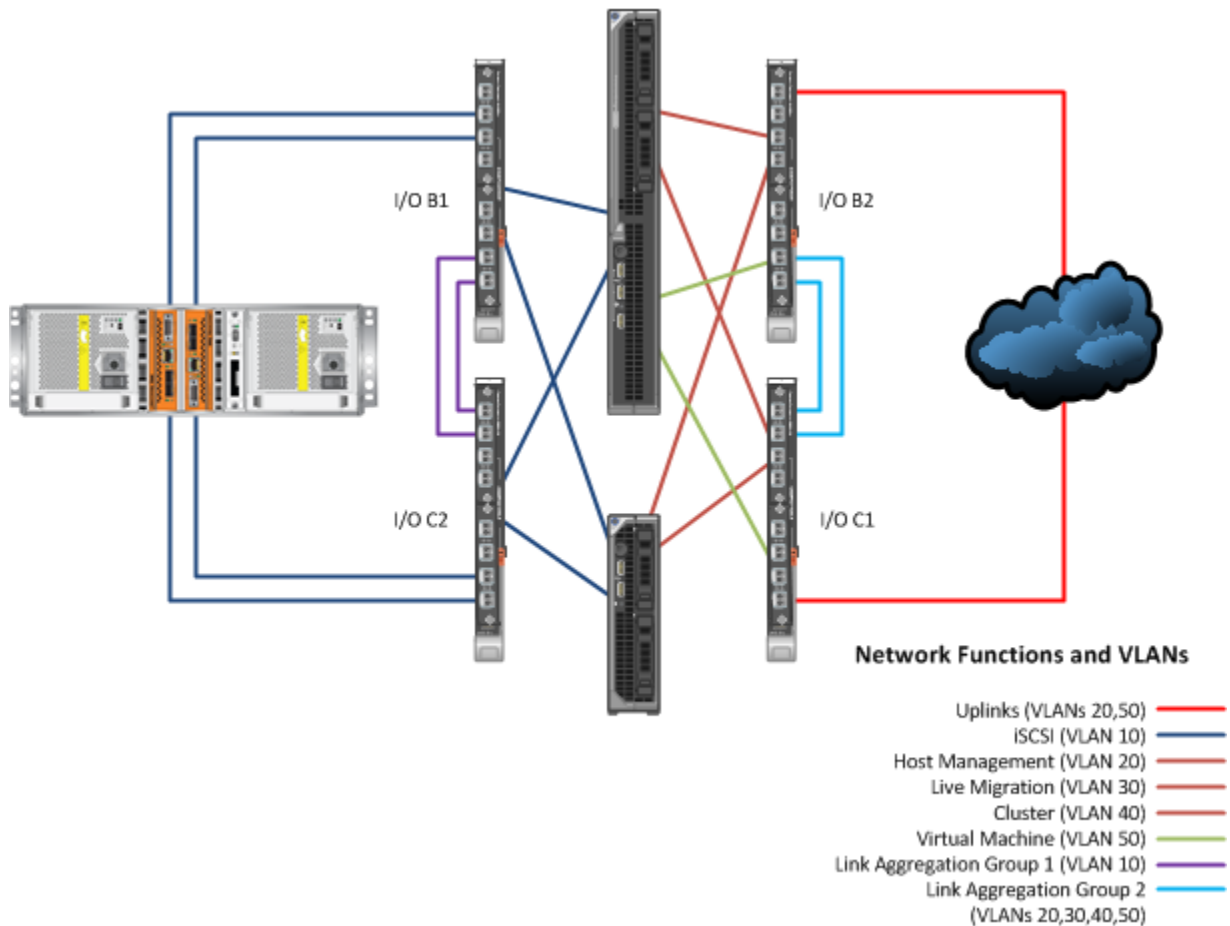
Network Architecture

The life blood of the configuration is a 10 Gbps Ethernet, powered by Dell PowerConnect blade switches. These managed, layer three, 10 Gbps Ethernet switches offer the enterprise-class level performance required for this configuration.

The switches are capable of supporting all configuration networks with the use of VLANs for traffic isolation, link aggregation, and network teaming for redundancy. As mentioned above and as is shown in Figure 2, reference configuration network traffic is broken down by function and assigned to a function-based VLAN and/or switch.

The switch configuration described in the appendix of this whitepaper (Appendix A: Reference Network Configuration) was used during the reference configuration development. Although modifications are possible, this switch configuration serves as the baseline configuration for the white paper since it will support all reference configuration hardware (up to six Hyper-V host servers plus two Hyper-V management servers).

Figure 2 - Network Fabric Overview



Data Networks, Subnets, and VLANs

In keeping with Microsoft best-practice recommendations, Ethernet traffic is broken down into separate networks; each network type is assigned a different network ID, switch, or subnet. Subnets are used to break the configuration network into smaller, more efficient networks. This also prevents excessive Ethernet packet collision, and provides a level of security by using switched-based isolation. Further isolation of network traffic types is achieved by using VLANs and physical isolation by dedicated switches.

As shown above in the Figure 2, the configuration defines the different network traffic types:

iSCSI Network (VLAN ID 10) is used for iSCSI communication between Hyper-V R2 hosts and EqualLogic PS6010X storage array(s). The two network switches in the B1 and C2 fabrics are dedicated for iSCSI communication. No VLANs are needed to isolate this traffic.

Host Management Network (VLAN ID 20) is used by cluster nodes for public or external communication. This network is also used by the existing infrastructure as well as the SCVMM R2 management server to connect to the cluster for management purposes.

Cluster Private Network (VLAN ID 30) is used by cluster nodes for inter-node communication. This network is also used by the server nodes in the failover cluster for cluster shared volumes.

Live Migration Network (VLAN ID 40) is used for the live migration traffic between Hyper-V R2 hosts in the failover cluster.

Virtual Machine Network (VLAN ID 50) is used for VM connectivity to the rest of the network. In the reference configuration, the VM network is isolated from all other traffic. Depending upon existing customer architecture or role of VMs to be deployed, the VM network may be shared with the Hyper-V Host Management Network.

Table 3 - Sample VLANs and Subnets

Traffic Type	Sample VLAN	Sample Subnet
iSCSI	10	10.10.10.0/24
Host Management	20	10.10.20.0/24
Cluster	30	10.10.30.0/24
Live Migration	40	10.10.40.0/24
Virtual Machine	50	10.10.50.0/24

Switch Trunking

The configuration uses two Link Aggregation Groups for Inter Switch Link (ISL) trunking. The trunk between the iSCSI switches allows all traffic and simply provides a highly available, high bandwidth connection between switches. The trunk used between the two switches supporting the host, cluster, live migration, and VM traffic are configured to allow VLANs 20,30,40, and 50. Similar to the previous trunk, this provides a high bandwidth and fault tolerant connection.

Spanning Tree

To prevent paths from incorrectly being blocked by a switching loop, Multiple Spanning Tree Protocol (MSTP) is used. MSTP is a per-VLAN STP and will ensure that a switch will block paths based on individual VLANs. If STP were used rather than MSTP the switch path for non-uplinked VLANs could potentially be blocked and would prevent cluster or live migration traffic.

Virtual Machine Network Fault Tolerance

Virtualization offers the ability to consolidate computing resources to improve management capabilities, improve return on investment of hardware resources, and increase network service availability metrics. Although these qualities are important, they often drive new design requirements that are not present in purely physical environments. One such consideration is the availability of network connectivity provided to the multiple VMs on a Hyper-V server. For example, in a non-virtualized environment, a physical network adapter carries the traffic for only one host. A loss of any networking component along the data path will only impact one physical server. In a virtual environment, that same physical networking adapter is carrying the traffic for all the VMs running on the host at that time (tens of VMs). This places an additional burden on the design of the VM network, and highlights the need to reduce the loss of connectivity due to network data path hardware failures.

To lessen this burden, network interface teaming is used to provide fault tolerant access for the hosts to the host management, cluster, live migration, and VM networks. Furthermore, the teaming software provides the interface for configuring the VLAN segregation. All of the configurations use the Broadcom Advanced Control Suite (BACS) 3.0 as the teaming configuration utility.

Uplinks (Off-Rack Connectivity)

Off-rack connectivity is provided by using the 10 Gbps Ethernet PowerConnect blade switches' external ports. These switches, when connected to a core switching infrastructure, provide extensive connectivity into the existing environment. The uplinks of the reference configuration allow host management and VM traffic (VLANs 20 and 50). If VMs and hosts are to share the same network, this may be consolidated to the single VLAN that hosts the traffic.

NOTE: The sample switch configuration shown in Appendix A: Reference Network Configuration includes configuration of two PowerConnect switch ports as uplink ports and utilize MSTP to optimize switching Paths.

M910 Network Configuration

The Hyper-V host servers require multiple Ethernet ports and advanced configuration provided by Broadcom software. Specifically, each server uses six Ethernet ports.

The following Figure 3 and Table 4 show the functional breakdown (network connections) of these ports and their mapping to the M8024 switches:

NOTE: To prevent cluster communication on interfaces not intended for use, all unused interfaces should be disabled.

Table 4 - M910 Switch Port Mapping

PowerConnect Internal Port Mapping	B1 - M8024 Mezz B _a NIC 0 Mezz B _b NIC 0	B2 - M8024 Mezz B _a NIC 1 Mezz B _b NIC 1	C1 - M8024 Mezz C _a NIC 0 Mezz C _b NIC 0	C2 - M8024 Mezz C _a NIC 1 Mezz C _b NIC 1
Blade Slot #2/#10	1/xg2, 1x/g10	1/xg2, 1x/g10	1/xg2, 1x/g10	1/xg2, 1x/g10
Blade Slot #3/#11	1/xg3, 1x/g11	1/xg3, 1x/g11	1/xg3, 1x/g11	1/xg3, 1x/g11
Blade Slot #4/#12	1/xg4, 1x/g12	1/xg4, 1x/g12	1/xg4, 1x/g12	1/xg4, 1x/g12
Blade Slot #5/#13	1/xg5, 1x/g13	1/xg5, 1x/g13	1/xg5, 1x/g13	1/xg5, 1x/g13

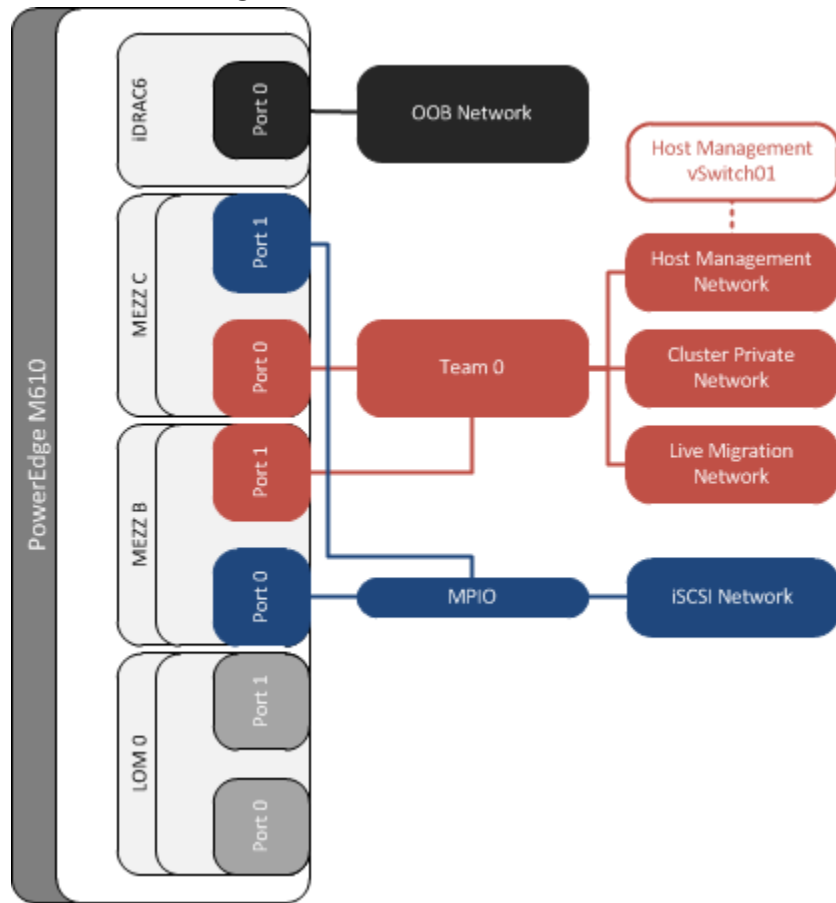
NOTE: The M910 servers have two Mezz B and two Mezz C risers. They are noted here and named on the server as Mezz Ba, Mezz Bb, Mezz Ca, Mezz Cb. The Ba and Ca interfaces map to the primary slot the server occupies while Bb and Cb map to the secondary slot that server occupies.

M610 Network Configuration

The Management servers require multiple Ethernet ports, and are coupled with an advanced configuration provided by Broadcom software. Specifically, each management server uses four Ethernet ports.

The following Figure 4 and Table 5 show the functional breakdown (network connections) of these ports and their mapping to the M8024 switches:

Figure 4 - M610 Network Connections



NOTE: To prevent cluster communication on interfaces not intended for use, all unused interfaces should be disabled.

Consult the *Hyper-V Networking Guide* for more information on the architecture of Hyper-V's virtual network stack, as well as a discussion of deployment considerations and procedures. Additionally, this guide will aid in the mapping of physical Ethernet ports to Windows-enumerated network adapters, the configuration of virtual switches, and the optimization of Windows server to work with more than one Ethernet connection.

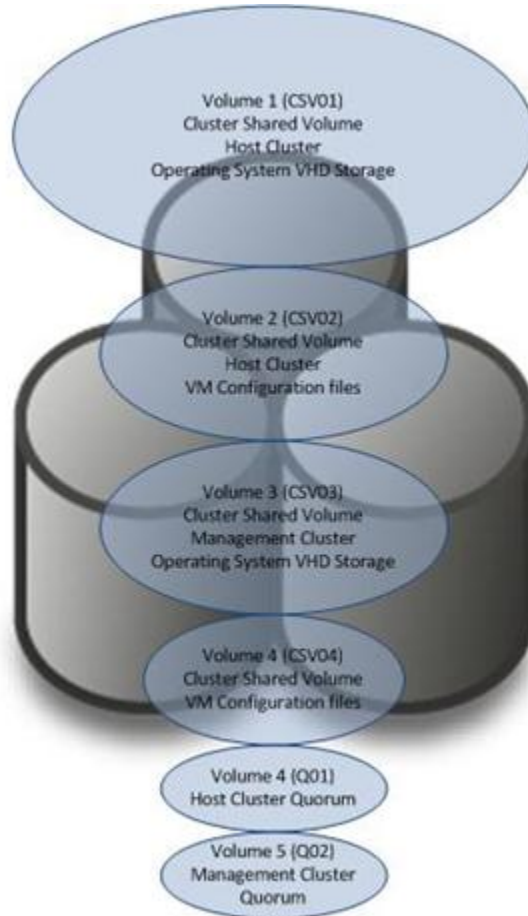
Table 5 - M610 Switch Port Mappings

PowerConnect Internal Port Mapping	B1 - M8024 Mezz B NIC 0	B2 - M8024 Mezz B NIC 1	C1 - M8024 Mezz C NIC 0	C2 - M8024 Mezz C NIC 1
Blade Slot #1	1/xg1	1/xg1	1/xg1	1/xg1
Blade Slot #9	1/xg9	1/xg9	1/xg9	1/xg9

Storage Architecture

The shared storage used in the reference configuration is highly flexible, configurable, and expandable. The reference configuration uses one volume per cluster for VM operating system storage, one volume per cluster for VM configuration files, and one volume per cluster for a quorum device. The distribution of volumes and data stored on the volume can be modified to best distribute I/O workloads (for example, one CSV for VM operating systems, one for VM configurations, one for random read and writes) This configuration is highly variable based on workload needs and customer specific requirements.

Figure 5 - Storage Volume Architecture



Shared-Storage Volume Size Considerations

Before the Dell EqualLogic PS6010X can be provisioned to servers, the array must be divided into storage pools. These pools are made up of physical hard drives with a certain type of RAID configuration. Before they can be used, these pools must be further divided into volumes, and must be assigned to a storage group. Volumes must be sized to accommodate not only the virtual machine hard drive (VHD), but also the size of the VM virtual memory and enough space for any VM snapshots.

NOTE: By default, when using Cluster Shared Volumes (CSV), VM snapshots are stored within the VM directory on the CSV of the shared storage. Make sure that enough space is available as the snapshots are created.

The following equation can be used as a starting guideline for sizing shared storage for virtual disks:

$VHD\ Volume = Size\ of\ all\ VHDs + (Space\ for\ .AVHD\ files)$

$VM\ Volume = 1.05 * (Allocated\ VM\ RAM + (Space\ for\ snapshots))$

VM snapshots record differences in VHD data between the previous snapshots and the base VHD file. Snapshots with few changes are roughly 100 MB at minimum. **Running out of shared storage causes the VM to pause and not restart until more disk space is made available.** The equation above suggests an extra 5% of disk space be allocated to prevent this error.

Array RAID Configuration

The storage array RAID configuration is highly dependent on the workload in your virtual environment. The storage array supports the following RAID types:

- RAID 5 generally provides more usable storage, but has less performance than RAID 10 in random I/O situations and requires additional overhead if a drive fails. It also provides the highest storage capacity at the expense of slightly lower performance and availability.
- RAID 6 will allow for the simultaneous failure of two drives resulting in an increase in availability at the expense of two drives capacity lost to parity.
- RAID 10 provides the best performance at the expense of storage capacity.
- Raid 50 provides a balance of performance, availability, and capacity.

The EqualLogic PS6010 series provides automatic storage tiering; therefore, a mix of RAID levels may be appropriate based on the application performance needs running inside the VM. Based on the RAID level chosen, the number of disks could increase or decrease as required to support the number of VMs configured.

The reference architecture validation uses RAID 10 with disk groups of 14 physical hard drives plus two hot spares; this configuration provides ~3.6TB per member of storage space to later be divided into smaller volumes. There is however, no single recommended configuration based on the RAID and volume mapping; consider the type of application and or datacenter requirements to determine what is appropriate for your environment and needs.

Jumbo Frames

Jumbo frames are enabled on the iSCSI network interfaces of all hosts to improve throughput. Jumbo frame size is set to 9600 bytes on all adapters communicating with the iSCSI array. This is also set on the iSCSI network switch. Jumbo frames are enabled by default and no configuration is required on PS6010 arrays.

Scalability

As workloads increase, the reference configuration can be scaled to provide additional resources. This section provides guidelines for adding blade servers, blade chassis, and storage arrays to the reference configuration.

Scaling Compute Resources

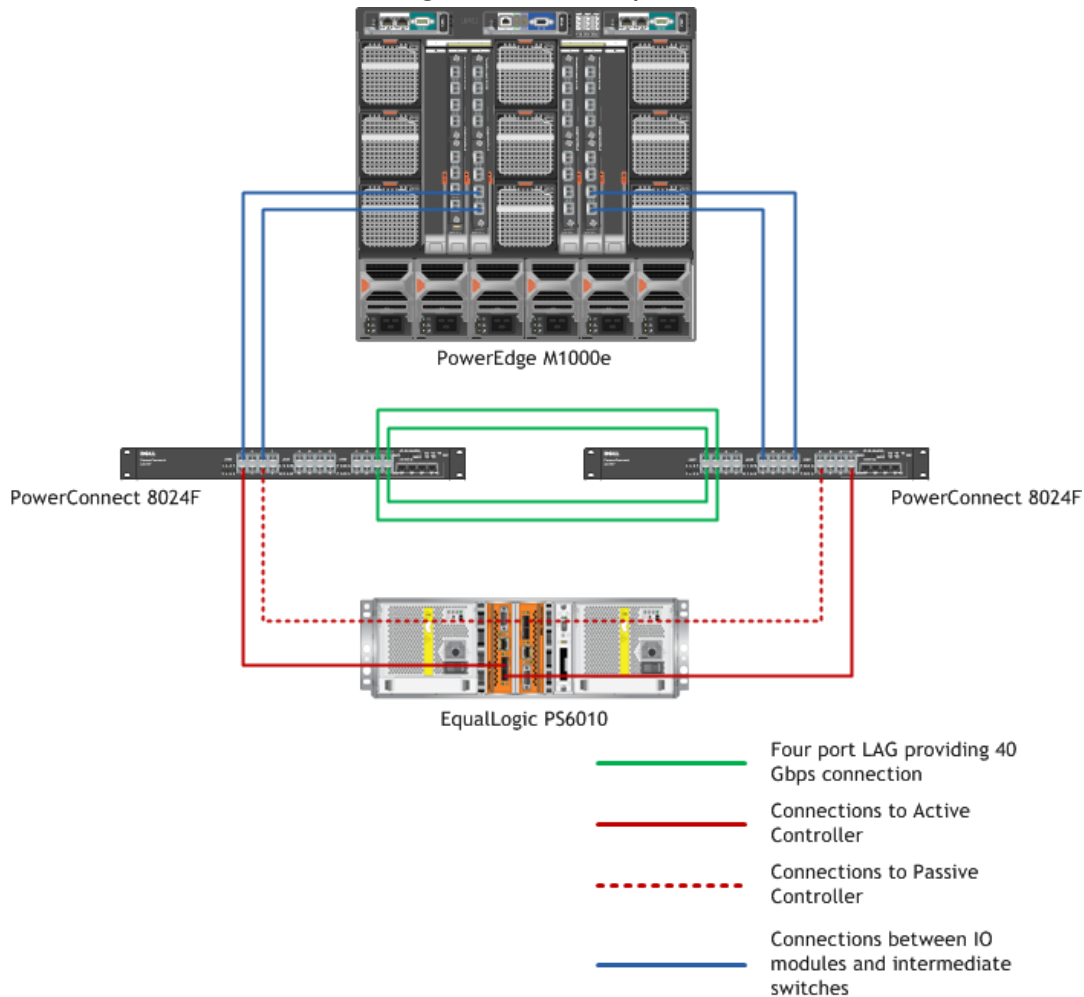
Adding blade servers: If the cluster has reached or is about to exceed its acceptable maximum resource utilization, additional blades can be added to the cluster.

Adding a blade chassis: If the chassis is full of blades and more compute resources are needed, an additional blade chassis can be added to the solution. Each new chassis will be uplinked to the external network infrastructure.

Scaling Storage Resources

The default solution can have up to three EqualLogic PS 6010 series arrays connected directly to the chassis. To scale beyond three PS6010 arrays per chassis, intermediate switches are required in the iSCSI fabric. Two PowerConnect 8024F switches can be used for intermediate switches, as shown in Figure 6. PowerConnect 8024F switches must be connected through Inter Switch Links (ISL) using LAGs. PowerConnect I/O module switches should not have ISLs, since this would create a loop. M8024 switches are connected to 8024F switches using links also configured in Link Aggregation Groups as shown in Figure 6.

Figure 6 - Scalable Option



Management

This section discusses several reference configuration management tools. Each subsection represents a different tool that is required for overall solution management. We start with a discussion of low-level hardware management options, and end with VM management; this is the order in which the reference configuration should be deployed.

Physical Server Management

The reference configuration is made up of several Dell PowerEdge servers. Although each of the servers plays a different role in the reference configuration, each can be managed in the same manner using both standard and optional server capabilities.

(a) Chassis Management Controller (CMC)

The Blade chassis is managed via the Chassis Management Controller (CMC). It is accessed through the physical VGA and USB ports, integrated KVM, Secure Shell (SSH), or via a web interface. Through the interface physical aspects of the chassis, I/O fabric, and servers can be managed.

(b) Keyboard, Video, Mouse (KVM)

Each blade chassis ships with integrated keyboard, video, and mouse interfaces that allows for a single entry point for console management of the reference configuration.

(c) iDRAC

The Integrated Dell Remote Access Controller6 (iDRAC6) is a systems management hardware and software solution that provides remote management capabilities, crashed system recovery, and power control functions for Dell PowerEdge systems. The iDRAC6 uses an integrated system-on-chip microprocessor to remotely monitor and control systems. The iDRAC6 co-exists on the blade chassis as part of the CMC. The server OS executes applications, while the iDRAC6 monitors and manages the server environment and state outside of the OS. Refer to the appropriate PowerEdge server manual at support.dell.com for more information about the iDRAC use and configuration options.

Storage Management

(a) EqualLogic Group Manager

The Dell EqualLogic PS6010X features three easy-to-use group management interfaces. The array can be accessed via serial console, telnet, or a user-friendly web interface. These interfaces enable complete setup and management of the storage array from a single interface.

(b) SAN HeadQuarters

SAN HQ from EqualLogic is an event and performance monitoring utility. It is capable of monitoring multiple SAN groups and members making a single site for easy monitoring of IOPS, throughput, capacity, and more.

SAN HQ is free of charge and available from:

<https://www.equallogic.com/support/download.aspx?id=8495> (registration required)

Systems Management

(a) Dell OpenManage

A web tool, Dell OpenManage Server Administrator (OMSA) can help customers manage individual servers and their internal storage from virtually anywhere at any time. OMSA is available with a Dell server at no additional charge. The tool:

- Helps to simplify single server monitoring, with a secure command line or Web-based management graphical user interface.
- Provides views of system configuration, health, and performance.
- Provides online diagnostics to help isolate problems, or shut down and restart the server (see <http://www.dell.com/openmanage> for more information about Dell OpenManage).
- Provides an integration point for external monitoring solutions such as Microsoft System Center Operations Manager 2007 R2 (SCOM).

(b) Unified Server Configurator (USC)

Dell servers ship with USC which 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. Refer to the server manuals at <http://support.dell.com>, or http://www.dell.com/content/topics/global.aspx/services/adi/sys_mgmt?c=us&l=en&cs=555 for more information about these and other Dell Management tools.

Virtual Machine Management

Microsoft Windows 2008 R2 with Hyper-V has the ability to locally manage Hyper-V servers, with limited cross-server manageability (power on, shut down, and migrate) using the Hyper-V Manager snap-in. For single server Hyper-V deployments, these functions might be sufficient. If multi-server management, rapid VM deployment, and many-to-many administration is required, you can use Microsoft add-on management products such as SCVMM R2. The failover cluster feature provides a single place to manage most of the critical functions of a virtualization environment. Additionally, Microsoft offers management products to aid with health monitoring and life cycle management as part of the System Center suite. More information can be found at <http://www.microsoft.com/systemcenter>.

Deployment

This section describes the deployment of the reference configuration. The steps outlined here do not address all possible configurations; rather, they serve as a baseline.

Network Configuration

The following section describes the setup of the reference configuration blade network switches.

To configure the Dell PowerConnect M8024 see the *Dell PowerConnect M8024 User's Guide* at <http://support.dell.com/support/edocs/network/M8024/en/index.htm> for more information.

See Appendix A: Reference Network Configuration for more information on the configuration used in the validation process.

1. Perform the initial configuration (for example, assign IP address)
2. Confirm that the Dell PowerConnect M8024 switches are running the latest firmware available from <http://support.dell.com>.
3. Configure the LAG between the two switches.
4. iSCSI Switches - Fabric B1 and C2 (PowerConnect M8024)
 - a. Configure the MTU for all ports to enable jumbo frames.
 - b. Disable Spanning Tree on all Hyper-V host and EqualLogic controller ports.
5. Host, Cluster, Live Migration, and Virtual Machine Switches - Fabric B2 and C1
 - a. Create the appropriate VLANs.
 - b. Allow VLANs on corresponding host ports.
 - c. Configure the MSTP regions as appropriate for your network.

Server Configuration

This section covers the different types of server setups used in the reference solution. Since Dell preinstalls the operating system that task is not covered, and the setup process begins after the OS activation and registration process is completed. The management servers (SQL and SCVMM R2), being VMs, are configured after the physical servers have been fully configured and the management cluster has been setup and are covered in Section 0.

All Microsoft® Windows 2008 R2 Hyper-V™ Servers

1. Enable server BIOS virtualization technology; for details on making changes to the BIOS, consult the Dell PowerEdge server documentation at <http://support.dell.com>. All default power and performance BIOS settings are used.

2. Using Microsoft Windows Update, install any needed security or program updates. Also ensure that all EqualLogic required updates are applied as described in release notes: https://www.equallogic.com/support/download_file.aspx?id=987.
3. Run the SUU utility that came with the Dell PowerEdge server. Updated software can also be downloaded from <http://support.dell.com>. When running the SUU, it is best to update all hardware component devices on the server to the latest firmware and driver versions available.
4. Install the Broadcom Advanced Control Suite and BASP features from Programs and Features.
5. Install the HIT Toolkit available at <https://www.equallogic.com/support/download.aspx?id=9821> (registration required). This will enable iSCSI initiator service and install the MPIO Device Specific Module for EqualLogic as well as install the Auto-Snapshot Manger/Microsoft Edition and User's Guide.
6. Join this server to the domain.
7. Install the Hyper-V role. After the install has completed, check for any updates to the Hyper-V role available from Microsoft Windows Update.
8. Install the Failover Cluster feature. (See the Failover Clustering Configuration section for details on cluster setup and configuration)
9. Create teams and assign appropriate IP addresses.
10. Enable Jumbo Frames on iSCSI network interfaces.

Storage Configuration

After servers have been initially configured, storage must be configured and volumes presented to the servers. This is performed in the following high level steps:

1. Perform the initial configuration of the PS6010X storage array (assign IP addresses to the network interfaces, and the group).
2. Create RAID groups and volumes.
3. Connect each host to the volumes assigned to their respective cluster (CSVs and Quorum). Ensure that MPIO is enabled when connecting to provide fault tolerance in the storage paths.
4. Initialize and create file system on each volume.

Failover Clustering Configuration

After the installation of the failover clustering feature on both Hyper-V management nodes and at least two of the Hyper-V R2 server nodes, you will need to complete the steps below to configure the high availability functionality. These steps are high-level configurations and must be completed for both the management and host clusters. These might not apply to your needs or environment.

1. Make sure the following items are completed:
 - a. All appropriate network ports are active and configured with their proper static assigned IP address (disable any unused network ports such as 1 Gbps Ethernet LOMs).
 - b. All nodes in the cluster have access to the shared storage volumes.
 - c. All nodes in the cluster have the same patch/hot fixes applied, as well as driver versions.
 - d. Run the Cluster Validation Wizard. Consult the Microsoft documentation for further details at [http://technet.microsoft.com/en-us/library/cc732035\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/cc732035(WS.10).aspx).

NOTE: Cluster Validation will fail with 802.1Q VLAN Teaming. Microsoft is aware of the issue and has noted that this is not a failure but just a warning and it will be changed to reflect that in a future release.

<http://support.microsoft.com/kb/974264>

2. Configure the Cluster networks; select names that are meaningful to the type of network traffic.
3. Add Cluster Shared Volumes.
4. Configure cluster quorum settings.
5. Create VMs.
6. Set the Network for Live Migration in the properties of the VM.

Management Configuration

After the clusters have been configured, the management virtual machines can be provisioned and deployed. These high level steps cover the deployment of SCVMM R2.

1. Deploy two VMs with resources appropriate for SQL and SCVMM R2 in the management cluster.
2. Install SQL Server 2008 and SQL 2008 SP1 on the SQL server.
3. Install SQL 2008 management tools on the SCVMM R2 server.
4. Install SCVMM R2.
5. Apply Microsoft updates for SCVMM.
6. Add hosts to the SCVMM Administration console.

How to Order the Reference Configuration

Dell simplifies the needs of IT customers by offering white papers, reference architectures, and video blogs to aid in the selection, deployment, and maintenance of solutions. In keeping with that methodology, it is Dell's intention to simplify ordering this reference configuration, and to enable configuration as needed to fix the needs of our customers.

This reference configuration can be ordered by using one of the two methods listed below:

- By building the reference configuration from scratch using the information in the Overview section.

By ordering a preconfigured bill of materials. The Solution ID listed in

- Table 1 can be used to select each of the reference configuration variations. This ID will greatly simplify the discussion with a Dell sales representative. The following is a link to all of Dell's business-ready configurations:

<http://content.dell.com/us/en/business/virtualization-business-ready-configurations.aspx>

Dell Global Services

Dell Global Services helps customers find suitable virtualization options to reduce the total cost of ownership, speed time to ROI, increase agility, and reclaim IT resources. Today's financial environment dictates that IT organizations reduce costs, 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 common virtualization implementation bottlenecks. Our methodology, tools, and processes are designed to speed up the implementation, ease migration scheduling, automate reporting, and provide process transparency. Dell Global Services also offers end-to-end solutions with a single point-of contact for hardware, software, services, and on-going support. Dell Global Services are focused on IT infrastructure services excellence.

To engage with Dell Global Services, refer to the following Web site:

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

Additional Reading and Resources

Dell PowerEdge Server Documentation and Hardware/Software Updates

For Drivers and other downloads: Visit <http://support.dell.com>, select “Drivers & Downloads”, enter a server service tag or select the server model and operating system version.

Dell PowerConnect Switch Documentation and Firmware Updates

Visit <http://support.dell.com>, select “Drivers & Downloads”, enter a server service tag or select the switch model and operating system version.

Dell Virtualization Documentation

Dell TechCenter Hyper-V R2 Blogs/Wiki Site:

<http://www.delltechcenter.com/page/Microsoft+Virtualization+-+Windows+2008+Hyper-V+R2>

Microsoft® Windows Server® 2008 R2 With Hyper-V™ for Dell™ PowerEdge™ Systems Important Information Guide:

http://support.dell.com/support/edocs/software/win2008/WS08_R2/en/index.htm

Business-Ready Configurations:

<http://www.dell.com/virtualization/businessready>

Microsoft® Hyper-V Documentation

Hyper-V Getting Started Guide:

[http://technet.microsoft.com/en-us/library/cc732470\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/cc732470(WS.10).aspx)

Hyper-V Update List for Windows Server 2008 R2:

<http://technet.microsoft.com/en-us/library/ff394763%28WS.10%29.aspx>

Supported Guest Operating Systems:

<http://www.microsoft.com/windowsserver2008/en/us/hyperv-supported-guest-os.aspx>

Hyper-V: Using Hyper-V and Failover Clustering:

[http://technet.microsoft.com/en-us/library/cc732181\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/cc732181(WS.10).aspx)

Microsoft® Management Software

Microsoft System Center Virtual Machine Manager (SCVMM):

<http://www.microsoft.com/systemcenter/virtualmachinemanager>

SCVMM 2008 R2 Deployment:

<http://technet.microsoft.com/en-us/library/cc764339.aspx>

Broadcom Networking

Broadcom Driver FAQ:

http://www.broadcom.com/support/Ethernet_nic/faq_drivers.php

Microsoft® Roles and Features

Introducing Windows Server 2008 Failover Clustering:

<http://technet.microsoft.com/en-us/magazine/2008.07.failover.aspx>

Appendix A: Reference Network Configuration

This appendix contains the network configurations for the Dell PowerConnect M8024 switches used in this configuration. The sample switch configuration will support six Hyper-V host servers, two Hyper-V management servers, three PS6010 arrays, and uplinks to an out-of-chassis switch.

This file includes the setup and configuration options used in the qualification of this configuration. Customers are encouraged to use this file as a baseline for modification to the switch setup and over all configuration architecture. The manuals at <http://support.dell.com> for the PowerConnect family of switches will aid with modifications to this baseline file.

For more information on Dell PowerConnect switches, see <http://www.dell.com/powerconnect>.

PowerConnect M8024 Switch 1 Configuration File Used for Host Management, Cluster, Live Migration, and Virtual Machine Traffic

```
configure

vlan database

vlan 20,30,40,50

exit

hostname "sw1"

stack

member 1 1

exit

ip address none

interface vlan 20

name "host management"

exit

interface vlan 30

name "cluster private"
```

```
exit
interface vlan 40
name "live migration"
exit
interface vlan 50
name "virtual machines"
exit
spanning-tree mode mstp
spanning-tree mst 1 priority 0
spanning-tree mst configuration
instance 1 add vlan 20
exit
spanning-tree mst configuration
instance 2 add vlan 50
exit
spanning-tree mst configuration
instance 3 add vlan 30
instance 3 add vlan 40
exit
spanning-tree mst configuration
name "Region1"
exit
interface range ethernet 1/xg1-1/xg24
switchport mode general
switchport general allowed vlan add 20,30,40,50 tagged
exit
interface ethernet 1/xg19
```

```
channel-group 1 mode auto

exit

interface ethernet 1/xg21

channel-group 1 mode auto

exit

interface port-channel 1

switchport mode general

switchport general allowed vlan add 20,30,40,50 tagged

exit

exit
```

PowerConnect M8024 Switch 2 Configuration File Used for Host Management, Cluster, Live Migration, and Virtual Machine Traffic

```
configure

vlan database

vlan 20,30,40,50

exit

hostname "sw2"

stack

member 1 1

exit

ip address none

interface vlan 20

name "host management"

exit

interface vlan 30

name "cluster private"

exit

interface vlan 40
```

```
name "live migration"

exit

interface vlan 50

name "virtual machines"

exit

spanning-tree mode mstp

spanning-tree mst configuration

instance 1 add vlan 20

exit

spanning-tree mst configuration

instance 2 add vlan 50

exit

spanning-tree mst 3 priority 0

spanning-tree mst configuration

instance 3 add vlan 30

instance 3 add vlan 40

exit

spanning-tree mst configuration

name "Region1"

exit

interface range ethernet 1/xg1-1/xg24

switchport mode general

switchport general allowed vlan add 20,30,40,50 tagged

exit

interface ethernet 1/xg19

channel-group 1 mode auto

exit

interface ethernet 1/xg21
```

```
channel-group 1 mode auto

exit

interface port-channel 1

switchport mode general

switchport general allowed vlan add 20,30,40,50 tagged

exit

exit
```

PowerConnect M8024 Switch Configurations Used for iSCSI Network

```
configure

stack

member 1 1

exit

ip address none

interface range ethernet 1/xg1-1/xg24

spanning-tree disable

mtu 9216

exit

interface ethernet 1/xg19

channel-group 1 mode auto

interface ethernet 1/xg21

channel-group 1 mode auto

exit

interface port-channel 1

switchport mode general

mtu 9216

exit

exit
```